National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

LARYNGEAL AIRWAY RESISTANCE AND PERCEPTUAL MEASURES IN PATIENTS WITH VOCAL NODULES

BY

MARNITA M. H. GRAMS ©

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND
RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE
IN SPEECH-LANGUAGE PATHOLOGY

DEPARTMENT OF SPEECH PATHOLOGY AND AUDIOLOGY

EDMONTON, ALBERTA
Spring, 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-10714-0

Canada

UNIVERSITY OF ALBERTA
LIBRARY RELEASE FORM

NAME OF AUTHOR: Marnita M.H. Grams

TITLE OF THESIS: Laryngeal Airway Resistance and Perceptual Measures in Patients with Vocal Nodules.

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: 1996

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Marnita Gram
Signature

#1706 10149 Saskatchewan Drive
Edmonton, AB T6E 6B6

Date: April 19/96.

THE UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **LARYNGEAL AIRWAY RESISTANCE AND PERCEPTUAL MEASURES IN PATIENTS WITH VOCAL NODULES** submitted by Marnita M.H. Grams in partial fulfillment of the requirements for the degree of Master of Science in Speech-Language Pathology.

Anne Putnam Rochet
Dr. Anne Putnam Rochet, Supervisor

Dr. Gary Holdgrafer

Harold Wiens
Professor Harold Wiens

april 19/96

ABSTRACT

This study utilized aeromechanical and perceptual tools to describe the changes that occurred in the voices of patients with vocal nodules who underwent voice therapy. A single-subject, A-B time-series design was used to track measures of laryngeal airway resistance in three subjects over a four-week baseline and an eight-week treatment period. Analysis of laryngeal airway resistance data consisted of visual inspection of graphic displays of resistance as well as the translaryngeal pressure and flow data from which resistance was calculated. Estimates of laryngeal airway resistance were correlated with perceptual ratings of voice samples (syllable trains and sentences) obtained from the subjects during the last week of baseline and each week of treatment. The Spearman Rank-Order Correlation Coefficient was used to correlate laryngeal airway resistance difference values and perceptual ratings completed by three clinicians in a paired-comparison paradigm.

Results suggested that laryngeal airway resistance did not change in the same direction or to the same degree with treatment across the three subjects, nor did the unique changes that occurred appear to be sensitive to changes in the status of subjects' vocal nodules with treatment. Significant correlations occured between data representing changes in laryngeal airway resistance with voice treatment and listeners' perceptions of those subjects' voices across the treatment period, but the validity of the correlations is rendered ambigious by a number of procedural, sampling, and psychophysical factors. Finally, intra- and inter-judge agreement data were variable and only moderately reliable which is consistent with reports in the literature for similar rating tasks and renders questionable the validity of the perceptual data used in the correlational analyses.

ACKNOWLEDGEMENTS

I would like to acknowledge my indebtness to my thesis advisor, Dr. Anne Putnam Rochet for her endless patience and advice in completing this project. Completion of this project would not have been possible without her generous support. I would also like to thank my other committee members, Dr. Phyllis Schneider, Dr. Gary Holdgrafer, and Professor Harold Wiens for their participation in the thesis process. A special thank-you to Dr. Holdgrafer for filling in at the last moment.

My thanks also go to clinicians Connie Zalmanowitz, Myka Osinchuk, and Lynette Jans at the Glenrose Rehabilitation Hospital for completing the perceptual rating task. I also wish to thank Connie for her help in recruiting patients and for being my mentor as a voice clinician. I am grateful to Dr. Glen Misko, otolaryngologist, for performing the nasendoscopy procedures. A special thank-you to my subjects for agreeing to participate in this study.

Last but not least I wish to express my deepest gratitude to my dear parents, Robert and Martha Grams for their support and encouragement throughout my graduate studies. Your love, encouragement, and prayers were my daily support. Thank-you for believing that I could accomplish the goals and dreams that I had set for myself.

#1706, 10149 Sask. Dr.
Edmonton, Alberta
Canada. T6E-6B6
Ph/FAX: (403) 439-8327.

March 22, 1996.

Barbara Kostyk Filipchuk
2009 Hen House Drive
Virginia Beach, Virginia.
23456
USA.

Re: Permission to reproduce a figure from your MSc. thesis.

I am writing to request permission to reproduce a schematic from your thesis entitled "Laryngeal Airway Resistance in Teachers with Vocal Fatigue".

Specifically, I request permission to reproduce your schematic which displays the Instrumental Array for Laryngeal Airway Resistance Data Collection in the thesis (unpublished) I have written to fulfill the requirements of my graduate degree program in speech-language pathology at the University of Alberta (Master of Science degree, Department of Speech Pathology and Audiology, University of Alberta. Supervisor: Dr. Anne Rochet).

Thank-you for your consideration of this request. A prompt reply to the address above would be most appreciated.

Sincerely,

Marnita Grams
Marnita Grams.

✓ Permission granted, with appropriate acknowledgement.

_______ Permission withheld.

____
Signature of copyright holder

96/04/09
Date

#1706, 10149 Sask. Dr.
Edmonton, Alberta
Canada T6E-6B6
Ph/FAX: (403) 439-8327

March 22, 1996.

Mr. Richard Lutz, President
Micro Tronics Corp
P.O. Box 16665
Chapel Hill, North Carolina
27516-6665
USA.

Re: Permission to reproduce two figures from the PERCI-PC software manual.

I am writing to request permission to reproduce two schematics from the PERCI-PC software manual (1986).

Specifically, I request permission to reproduce Figure 5B Laryngeal Resistance Configuration and Figure 7 from the above-mentioned manual in the thesis (unpublished) I have written to fulfill the requirements of my graduate degree program in speech-language pathology at the University of Alberta (Master of Science degree, Department of Speech Pathology and Audiology, University of Alberta. Supervisor: Dr. Anne Rochet).

Thank-you for your consideration of this request. A prompt reply to the address above would be most appreciated.

Sincerely,

Marnita Grams
Marnita Grams.

__✓__ Permission granted, with appropriate acknowledgement.

______ Permission withheld.

Signature of copyright holder

4/11/96
Date

I would appreciate receiving a copy of thesis. Thank you.
R Lutz

TABLE OF CONTENTS

Chapter Page

I INTRODUCTION 1

II REVIEW OF THE LITERATURE 3

  Vocal Nodules 3

  Laryngeal Airway Resistance 6

  Perceptual Measures of Voice Quality 20

  Relationships between Aerodynamic and Perceptual Measures 25

  Conclusion of Literature Review 28

  Statement of Problem 29

  Statement of Purpose 29

III METHODOLOGY 30

  Design 30

  Experimental Variables 30

  Experimental Phases 31

  Subjects 32

  Procedure 34

    Subject Selection 34

    Experimental Protocol 35

    Laryngeal Airway Resistance Data Collection and Measurement 36

    Acoustical Data Acquisition and Analog-to-Digital Conversion 43

    Perceptual Data Acquisition 44

  Data Analysis 49

IV RESULTS 51

  Flexible Fiberoptic Videonasendoscopy 51

Chapter Page

Laryngeal Airway Resistance 53
    Menstrual Cycle 65
    Colds 66

Perceptual Ratings 66
    Order Effect for Paired Sample Presentations 66
    Trends in Perceptual Data 68
    Intrajudge Agreement 73
    Interjudge Agreement 75

Laryngeal Airway Resistance Correlated with Perceptual Ratings 78

V DISCUSSION 86

Laryngeal Airway Resistance 86
    Instrumental and Procedural Errors 89
    Vocal Intensity 90
    Baseline Stability 91
    Menstrual Cycle 92
    Colds 92
    Summary of Correspondence between LAR and Treatment Outcome 93
Correlations Between Aeromechanical and Perceptual Measures 93
    Procedural Threats to the Validity of the Perceptual Data 96
    Psychophysical Threats to the Validity of Data used in Correlations 98
        The Nature of the Perceptual Sample 98
        Listener Reliability and the Nature of the Listening Task 99
    Summary of Correlations between LAR and Perceptual Data 102
External Validity Issues 102
Clinical Validity 103

Conclusions and Suggestions for Further Research 103

REFERENCES 106

APPENDICES 111

LIST OF TABLES

Tables Page

1. Subject Information....52

2. Intrajudge Order Effects....67

3. Intrajudge Percent Exact Agreement....74

4A. Percent Exact Agreement among Ratings for all Judges....76

4B. Percent Exact Agreement between Pairs of Raters....77

5. Correlations between LAR Differences and Perceptual Ratings....80

LIST OF FIGURES

Figures Page

1A. Placement of Mask and Catheter 39

1B. Instrumental Array for Laryngeal Airway Resistance Data Collection 40

1C. Laryngeal Airway Resistance Sample Data 42

2A. Instructions presented to Raters 47

2B. Response Screen Display for Raters 48

3A. Subject 1 Resistance 54

3B. Subject 1 Pressure 55

3C. Subject 1 Flow 56

4A. Subject 2 Resistance 57

4B. Subject 2 Pressure 58

4C. Subject 2 Flow 59

5A. Subject 3 Resistance 60

5B. Subject 3 Pressure 61

5C. Subject 3 Flow 62

6A. Subject 1 Syllable Trains Perceptual Trends 69

6B. Subject 1 Sentences Perceptual Trends 69

7A. Subject 2 Syllable Trains Perceptual Trends 70

7B. Subject 2 Sentences Perceptual Trends 70

8A. Subject 3 Syllable Trains Perceptual Trends 71

8B. Subject 3 Sentences Perceptual Trends 71

9. LAR Difference Values for Subject 1 83

10. LAR Difference Values for Subject 2 84

11. LAR Difference Values for Subject 3 85

LIST OF APPENDICES

Appendix Page

A. Subject's Consent Form allowing Primary Investigator to contact Subjects ... 111

B. Telephone Screening Form ... 112

C. Vocal Nodules Treatment Study Information ... 117

D. Subject's Participation Consent Form ... 119

E. Voice Case History ... 123

F. Voice Treatment Information for Subjects ... 128

G. Complete Perceptual Data Sets (G-1 to G-6) ... 131

CHAPTER I. INTRODUCTION

It has been proposed that the information provided by measures of laryngeal airway resistance, or the extent to which the larynx opposes airflow through it during phonation, can aid clinicians in understanding and evaluating disordered laryngeal function (Smitheran & Hixon, 1981). Smitheran and Hixon (1981) developed a non-invasive procedure for estimating laryngeal airway resistance during vowel production. Subsequent investigations, using variants of the Smitheran and Hixon procedure (Hoit & Hixon, 1992; Leeper & Graves, 1984; Lotz & Netsell, 1986; Melcon, Hoit & Hixon, 1989; Netsell, Lotz & Shaughnessy, 1984), have produced a modest data base of laryngeal airway resistance measures across a broad range of ages. Results of some of this work indicated that subjects with voice disorders fell outside the normal limits of laryngeal resistance on either subglottal pressure, air flow, or both (Netsell, Lotz & Shaughnessy, 1984).

Individuals in need of voice therapy may be distinguished on the basis of the perceptual quality of their voices. A patient with vocal nodules, for example, is often perceived as having a breathy, hoarse, low-pitched vocal quality. The breathiness may be related to the presence of the nodules, which prevent complete approximation of the vocal folds. The nodular mass accumulation at the midpoints of the vibrating vocal folds has the effect of increasing the mass on the cover of the vocal folds, thereby decreasing pitch (Boone & McFarlane, 1988), and perturbing vocal fold vibration thereby producing hoarseness (Colton & Casper, 1990). Voice therapy is aimed at decreasing the vocally abusive behaviors which have resulted in the formation of the nodules, and teaching patients how to achieve the optimal use of their vocal mechanisms. The reduction of vocally abusive behaviors and the provision of techniques for optimal voice use eventually result in a perceptible reduction of the hoarseness and breathiness in the vocal quality. Unfortunately such perceptual judgments are subjective. Furthermore "terminology problems have been duly noted in the literature" (Reed, 1980, p.157). What is perceived as "rough" by one listener may be perceived as "hoarse" by another. Thus there is a need for more objective data measurement regarding the efficacy of voice therapy (Reed, 1980).

Because estimates of laryngeal airway resistance in voice-disordered subjects have been reported to differ from those in individuals with normal voices, and because voice therapy results in changes in laryngeal function (i.e., better vocal fold approximation, less muscular tension) and laryngeal structure (i.e., reduction or disappearance of vocal nodules), it is logical to explore whether laryngeal airway measures might change across time as a dysphonic patient's vocal quality approaches normalcy as a result of treatment.

The purpose of this investigation was to obtain simultaneously acoustical and aeromechanical samples of voice in subjects with vocal nodules across pre-treatment baseline and treatment phases to determine if aeromechanical data are reliable indicators of a response to voice therapy and if a relationship exists between them and perceptual judgements of the corresponding acoustical data. A review of literature pertinent to this experimental endeavor follows and covers the following four areas: (1) the nature and acquisition of vocal nodules, (2) the nature and clinical utility of laryngeal airway resistance estimation, (3) perceptual studies related to problems of listener judgements of dysphonia, and (4) the combined use of aeromechanical and perceptual measures in describing and monitoring changes in dysphonia.

CHAPTER II. REVIEW OF THE LITERATURE

Vocal Nodules

Because subjects with dysphonia and vocal nodules were the source of the aeromechanical and perceptual data analyzed in this research, some information about the nature and etiology of nodules and their effects on voice production is warranted at the outset of this review. The average voice clinician is apt to encounter many patients with vocal nodules. Indeed Colton and Casper (1990) state that "vocal nodules are the most common benign lesions of the vocal folds in both children and adults" (p.61). Hyperfunctional use of the voice is viewed as a causative factor in the development of vocal nodules as well as of other organic pathologies of the larynx such as polyps, contact ulcers, vocal-fold thickening and vocal fold edema (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989). "The term 'vocal hyperfunction' refers to conditions of abuse and/or misuse of the vocal mechanism due to excessive and/or 'imbalanced' muscular forces" (Hillman et al., 1989, p. 373). "Musculoskeletal tension disorders" (Aronson, 1980) and "muscular tension dysphonia" (Morrison, Rammage, Gilles, Pullan, & Hamish, 1983), are examples of other terms that describe voice disorders related to hyperfunction (Hillman et al., 1989).

> "Vocal fold vibration results from an alternating balance between [subglottal] air pressure on the folds, which drives them apart, and elastic restoring forces, which draw them together.... The magnitudes of these forces depend on interactions among several factors; the amount of transglottal pressure drop, the separation of the vocal folds, and the tension on and the stiffness in the folds " (Hillman et al., 1989, p.374).

Certain conditions will produce "abnormalities in physiological states and acoustic source characteristics and some of these abnormal conditions may be referred to as vocal hyperfunction" (Hillman et al., p.374). Increased or poorly regulated laryngeal muscular tension is one such state that may result in vocal hyperfunction. Increased muscular tension produces abnormally stiff and tightly approximated vocal folds accompanied by increased subglottal air pressures and possibly larger amplitudes of vocal fold vibration. Such conditions produce abnormal glottal cycles in which the closing phase "is characterized by higher vocal-fold closing velocities and collision forces" (Hillman et al., p. 375). At first the increased muscular tension and collision forces may simply result in vocal fatigue, and the voice may be characterized as "strained". If these conditions prevail for a sufficient length of time, however, the excessive trauma of the vocal fold tissues results in inflammation and edema. Differences between the opposing folds in the mass and stiffness of their covers caused by edema may be responsible for inconsistencies of vocal fold approximations from cycle to cycle. (Hirano 1981b, cited in Hillman et al., 1989). If the vocal folds are not approximating one another completely over their entire length, there may be increased airflow through the glottis. At this point in the hyperfunctional behavior pattern, if the vocal abuse ceases or is significantly reduced, the edema will resolve and the dysphonia will likely disappear. However, if at this stage these hyperfunctional behaviors persist, then the vocal fold tissues undergo further changes, ultimately resulting in more substantial vocal fold lesions such as nodules, polyps and contact ulcers. It has been suggested that the site and extent of trauma induced as the vocal folds approximate (collide) with one another determine the specific type of vocal lesion that results (Hirano, Kuita, Matsuo & Nagata, 1980, cited in Hillman et al., 1989).

The structural development of vocal nodules is described by Colton and Casper (1990). Early nodules tend to be fairly soft and may be reddish in color. At this stage they are mostly vascular and edematous. In addition the rest of the vocal fold may also be edematous, and the entire larynx may appear slightly inflamed. "With continued trauma, the tissue undergoes hyalinization and fibrosis and becomes firm" (Colton & Casper, 1990, p.91). Thus chronic nodules tend to be hard, white, thick and fibrosed. Nodules in the chronic stage are usually bilateral but may not always be symmetrical in size. Nodules are generally located at the juncture between the anterior and middle thirds of each vocal fold, and the nodule site involves the membrane covering the vocal ligament as well as the vocalis inner border (Boone & McFarlane, 1988).

The presence of nodules has the effect of increasing the mass on the cover of the vocal folds and results in a decrease in pitch (Boone & McFarlane, 1988). According to Colton and Casper (1990), "extra mass at the midpoint of the vibrating vocal folds results in increased periodicity of vibration, greater frequency perturbation and greater hoarseness" (p.93). Boone and McFarlane note that "during phonation, when bilateral nodules approximate one another an open glottal chink usually results on each side of the nodule" (p.61). Such "chinking" occurs when the nodules approximate one another in exact juxtaposition. The glottal chink causes faulty approximation of the vocal folds and results in a breathy vocal quality (Boone & McFarlane, 1988).

Aerodynamic measures reported by Hillman et al. (1989) indicated that patients with nodules displayed excessively increased levels of transglottal pressure. The high levels of pressure presented by these patients likely result from the difficulties in producing voice "in the presence of heightened levels of laryngeal muscle activity that cause the vocal folds to be abnormally stiff and tense and increase the force of adduction" (p.382). Patients with nodules also exhibited abnormally high values of airflow (Hillman et al., 1989). Higher rates of transglottal airflow during phonation could be attributed to several phenomena associated with the presence of nodules on the vocal folds; incomplete closure of the glottis at the region of the nodules, a posterior glottal chink, and increased subglottal air pressure.

The nature of the lesions in vocal nodules, the behavioral factors underlying their development, and the quality of the dysphonia associated with their presence all suggest that information about translaryngeal pressures and flows and laryngeal airway resistance should be salient to an understanding of the effects of nodules on voice production. Phonation occurs at an inaccessible and extremely sensitive place in the airway, however, making direct measurements of transglottal pressure and flow invasive and uncomfortable in all subjects, and sometimes impossible in many. Fortunately a non-invasive, indirect alternative method of sampling this information is now available and was used in this research. That method will be described in detail in the text that follows.

Laryngeal Airway Resistance

Development of the Method.

The larynx functions as a mechanical valve between the lower and upper airways. Smitheran & Hixon (1981) suggest that much can be learned about the larynx's function "during vowel production from knowledge of the magnitude of coupling the larynx permits between the trachea and pharynx" (p.138). The extent to which the larynx opposes the airflow through it, termed "laryngeal airway resistance" (LAR) , is one source of such knowledge. Laryngeal airway resistance (R_{law}) cannot be measured directly but must be estimated from the ratio of translaryngeal pressure (P_{tl}) (the difference between tracheal and pharyngeal pressures) to translaryngeal flow (V_{tl}) (the flow through the larynx during phonation). This relationship is portrayed in the following formula: $R_{law} = P_{tl} / V_{tl}$. The use of laryngeal airway resistance measures has been tauted to have clinical value, helping clinicians to understand, evaluate and manage disordered laryngeal function. It was not until the last decade, however, that a noninvasive procedure was introduced by Smitheran and Hixon (1981) for estimating laryngeal airway resistance during the production of vowels. Before then, aeromechanical data sampling methods had not been explored that would enable clinicians to collect information on laryngeal airway resistance without the necessity of a tracheal puncture or an esophageal balloon to sample subglottal pressure. Such methods are disadvantageous for routine clinical use because they are invasive, cause discomfort for the subjects, require the assistance of a physician and are not without physical risks. Leeper and Graves (1984) maintained that voluntary compliance for such procedures on a routine basis is very difficult to obtain from adults, while compliance from children is untenable. Leeper and Graves also suggested that invasive methods did not consider the entire effect of total upper airway resistance on laryngeal activity during speech.

Thus Smitheran and Hixon's development of a simple, non-invasive method for estimating laryngeal airway resistance during vowel production created a "breakthrough" of sorts in the area of clinical laryngeal aeromechanics. As all other subsequent investigations in this area have based their rationale and procedures on those of Smitheran and Hixon, consideration of the conceptual development and salient features of their method is warranted.

The conceptual development of the method has origins in general fluid dynamics theory. Such a theory predicts that "in many fluid-filled systems, certain combinations of valving adjustments enable estimates of upstream pressure-flow events made at sites that are substantially downstream of the measurement sites of interest" (Smitheran & Hixon, 1981, p.138). When this theory is applied to the respiratory passages, certain combinations of valving adjustments of the larynx, velum, and oral structures enable researchers to estimate tracheal and pharyngeal pressures, and translaryngeal flow from pressure-flow events downstream of the trachea, pharynx, and larynx . Smitheran and Hixon (1981) reasoned that the

> "key to this combination is the production of an utterance which, when elicited from a subject, will call up and serially order various combinations of valving adjustment, such that it is possible to obtain discontinuous estimates of tracheal pressure, pharyngeal pressure, and translaryngeal flow associated with the vowel elements in the utterance from the simultaneous and continous recording of relevant down-stream pressure flow events" (p.139).

Utterance Design.

A simple utterance which would fulfill this requirement would be one in which repeatedly alternating voiceless stop-plosives and vowels would occur. Oral pressure during the voiceless stop consonant can be used to infer discontinuous estimates of tracheal pressure, whereas airway opening flow during the vowel can be utilized to obtain estimates of translaryngeal flow. An estimate of tracheal pressure can be obtained from measurement of oral pressure because during the production of voiceless stop plosives, the oral and tracheal pressures are identical. During the latter part of the closed phase of the voiceless plosive (when the upper airway is sealed tightly and the laryngeal airway is open) the pressure within the oral cavity behind the site of airway occlusion will equilibrate with that of the pressure within the trachea and lower airways (Smitheran & Hixon, 1981). The specific moment of oral-tracheal pressure equilibrium occurs with the peak oral pressure of the voiceless plosive (Netsell, 1969; Shipp, 1973). "Airway-opening flow can be used as an estimate of translaryngeal flow during vowel production because of the continuous airflow through the upper airway when the mouth is open" (Smitheran & Hixon, 1981, p.139). Any small flow generated by articulatory motion is considered zero if the airway opening measurements are made during a time when the airway is in a quasi-static state (such as during the middle portion of vowels) (Smitheran & Hixon, 1981).

The utterance chosen by Smitheran and Hixon (1981) was the syllable /pi/, repeated in three separate trains of seven productions each. For each syllable train, subjects were required to take a breath of twice normal depth and to produce the seven /pi/ syllables on a single continuous expiration at a rate of 1.5 syllables/s. Productions were to be made at normal loudness and habitual pitch, with equal stress on each syllable. Each of the seven vowels in the utterance train was to be prolonged slightly. Smitheran and Hixon maintained that their choice of utterance was a good one for the following reasons:

The /p/ consonant, being a voiceless plosive, results in oral and tracheal pressure equilibrium at the moment of peak oral pressure during the stop (Shipp, 1973 cited in Smitheran & Hixon, 1981) The /p/ phoneme is the earliest appearing voiceless stop-plosive in phoneme development (Prather, Hedrick & Kern, 1975) and thus can be used even with young children.

The place of /p/ production is the most anterior of all voiceless stop plosives in the English language. Thus "the measurement of oral pressure can be accomplished with the least likelihood of recording artifact from blockage of a probe by saliva, and with the least likelihood of recording artifact from articulatory interference with a probe" (Smitheran & Hixon, 1981, p.140) In addition, the place of /p/ production (at the lips) increases the likelihood that an airtight seal can be formed around the sensing probe. Finally, the labial place of production makes it easy for the experimenter or clinician to determine visually whether closure of the oral cavity is adequate. This becomes important in testing persons with structural or neuromuscular disorders.

The /i/ vowel (along with other high vowels) has been empirically proven to result in airtight velopharyngeal closure (Thompson & Hixon, 1979) Flow artifacts and air leakage around the flow-collection mask worn by subjects are minimized through the use of /i/ as this vowel does not involve extensive articulatory movements of the mandible or perioral region.

Breath grouping, stress patterning and rate were also taken into account in the design of the utterance. Attempts were made to control the form of the breath group used by the subjects. "By telling the subjects to take a breath of twice normal depth before utterance and to produce utterance on a single continous expiration, it was reasoned that breath group conditions would be similar to those of conversational speech" (Smitheran & Hixon, 1981, p.140). To closely approximate subjects' "normal" manners of using the larynx and to provide a natural routine, subjects are asked to perform the utterances with normal loudness and habitual pitch. The choice of a syllable as an utterance unit was prompted by the simplicity of subject performance and the ease with which investigators could monitor linguistic stress and utterance rate. It was desired that equal stress be placed on each syllable to maximize the consistency of the force provided to the larynx and the upper airway through the respiratory system (to maintain a fairly flat tracheal pressure contour). Selection of repetition rate of 1.5 syllables/s was also selected in order to ensure that the larynx assumes a quasi-static resistance and to ensure the maintenance of adequate velopharyngeal closure during the utterance. Thompson and Hixon (1979) found that utterance rates below 1.5 syllables /s are often accompanied by intermittent openings of the velopharynx.

Smitheran and Hixon's Results.

Smitheran and Hixon (1981) reported data for 15 adult male subjects, aged 21 to 40 years, to support the introduction of their method. Resistance values for this subject group averaged 35.7 cm H_2O /LPS (± 3.3 cm H_2O) and ranged from 30.0 to 43.1 cm H_2O/LPS. Smitheran and Hixon (1981) compared their data to the weighted mean data for three earlier investigations that used invasive methods (tracheal punctures) to sample subglottal pressure in similar sized groups of adult male subjects producing sustained vowels (Kunze, 1962; Sant & Logemann, 1970; Shipp & McGlone, 1971). The mean values for the two sets of data (pooled values for the earlier three invasive experiments and those of Smitheran & Hixon) differed by only 1.5 cm H_2O. This favorable comparison supported the reliability and validity of the non-invasive method of Smitheran and Hixon.

Critique of Smitheran and Hixon's Method by Rothenberg.

Rothenberg (1982) raised several questions regarding Smitheran and Hixon's use of oral pressure during the voiceless plosive /p/ as an estimate of the translaryngeal pressure associated with adjacent vowels. Two of his concerns will be presented here briefly because they are germane to the syllable utterance rate and the phonetic makeup of their speech sample.

The utterance rate chosen by Smitheran and Hixon (1.5 syllable/s) was influenced by the need to ensure airtight velopharyngeal closure which could have been compromised at utterance rates below 1.5 syllables/s. Rothenberg (1982) favored a syllable repetition rate somewhat higher than 1.5 syllables per second because of a concern that a subject might make separate respiratory gestures for each syllable at rates slower than about two per second. Smitheran and Hixon (1982) argued that their specification of a breath group beginning at twice resting tidal depth and a steady expiratory excursion constituted a "major precaution taken to avoid any problem with continuity" (p.220). These investigators cited pilot data measuring esophageal pressure and oral pressure simultaneously with kinematic recordings of the chest wall and found no evidence of discrete respiratory gestures. Thus Smitheran and Hixon concluded that their choice of a rate of 1.5 syllables/s was sound. They suggested that if rate were to be increased, individuals with neuromuscular involvement might be unable to perform the task adequately.

Rothenberg (1982) also raised the question of whether possible error could be introduced into translaryngeal pressure estimates during vowels by aspiration associated with preceding stop-plosive release. Rothenberg was concerned about the drop in tracheal pressure that occurs when the airway is momentarily open to the atmosphere and argued that longer periods of aspiration might change the magnitude of the drop in tracheal pressure. In designing their method, Smitheran and Hixon attempted to avoid such problems. In examination of data from their own as well as other laboratories (i.e., Netsell, 1969; Shipp, 1973) they made the observation that "tracheal pressure typically returned to its consonant pre-release value once the utterance had progressed well into the segment of the post-consonantal vowel" (p.221). This caused Smitheran and Hixon to make their flow measurements at the approximate mid point of the vowel. According to Netsell (1969), and Shipp (1973), the pre-release pressure on the /p/ is an excellent indicator of the subglottal pressure at the point of measurement that Smitheran and Hixon used during the following vowel.

Thus, while they acknowledged Rothenberg's suggestions and concerns, Smitheran and Hixon argued that their choices of utterance rate and speech sample were warranted and constituted "certain compromises that need to be made in transferring technological advances from the laboratory to the clinic" (1982, p.222).

The Method as a Routine Clinical Tool.

Smitheran and Hixon (1981) reasoned that the results of their study were encouraging and showed promise for the usefulness of their method as a routine clinical tool. There are several reasons to suggest that the method has practical value. First the data they reported have low inter-subject variability. Second, their results showed a high intrasubject reliability from analyses made on three performances of the utterances specified in the protocol. In addition, the utterance is simple, the subject experiences no discomfort during data collection, and the procedure is essentially non-invasive. As a result, a physician does not need to be physically present to supervise the data collection, and no special ethical considerations are required. Finally, the use of this method allows a wide variety of voice disorders to be studied. Smitheran and Hixon suggest that "such data can be used to document and chart changes in function, to make numerical comparisons between subject performance and baseline norms, to compare values from clinic to clinic, and to quantify the success of various mechanical management programs" (1981, p.144) (such as Teflon injections). Data obtained from laryngeal airway resistance can also complement voice assessment data obtained through other methods of observation such as perceptual judgement.

Smitheran and Hixon described the use of their method with two individuals in whom laryngeal airway resistance measures were sampled before and after a period of voice therapy. One individual had vocal nodules, and one exhibited dysarthria. The individual with vocal nodules exhibited a low-pitched, harsh, breathy voice, and indirect laryngoscopic examination revealed bilateral vocal nodules. Resistance was found to be 22.9 cm H_2O/LPS. Smitheran and Hixon concluded that the individual's low laryngeal airway resistance could be attributed to the presence of the vocal nodules which prevented firm approximation of the vocal folds along their length. After an intensive three month voice therapy program which involved identification and elimination of abusive laryngeal behaviors and voice modification procedures (such as raising vocal pitch), as well as biofeedback about the pressure and flow parameters of her laryngeal airway resistance, the subject was dismissed from therapy. Her vocal nodules were no longer present, and laryngeal airway resistance had increased to 34.3 cm H_2O/LPS. This subject represents the first individual on whom laryngeal airway resistance changes have been documented in response to a vocal management program.

Although Smitheran and Hixon reported the results of this case study as an example of the clinical utility of their method, the validity of the example may be questionable. First, they reported only one pre-treatment measure and one post-treatment measure of laryngeal airway resistance. In a single-case research design it is necessary that baseline measurements be collected over a longer period of time, until the data have stabilized. Only when baseline data have stabilized, can they be compared to data from a subsequent treatment phase. In addition, single-subject experimental designs require that data need to be sampled at multiple points throughout the treatment phase of a clinical study. The multiple probes in the treatment phase provide an opportunity to observe a trend indicating a treatment response, should one occur. Furthermore, when a single-subject research design is used in clinical research, it is recommended that the design be replicated on at least four individuals with similar clinical profiles (McReynolds & Kearns, 1983) and at least once within a single individual. Unfortunately, Smitheran and Hixon (1981) reported only pre- and post-treatment data on two individuals, each presenting a different disorder.

Its experimental flaws notwithstanding with respect to treatment measurement, the Smitheran and Hixon report (1981) "paved the way" for future clinical research using laryngeal airway resistance. Smitheran and Hixon suggested studying a variety of dysphonic subjects with their method to discover how resistance is influenced by disordered production of voice and to evaluate changes with management. It was their utimate hope that a normal range of laryngeal airway resistance could be established with cut-off points below and above which dysfunctional laryngeal airway resistance could be identified.

Other Studies of Laryngeal Airway Resistance

Since Smitheran and Hixon's presentation of a non-invasive clinical method for the estimation of laryngeal airway resistance in 1981, a small number of other studies have been completed that expand the data base on laryngeal airway resistance values to include both male and female speakers in various age groups. The results of these studies have contributed to a modest data base on laryngeal airway resistance and will be summarized below. These studies have also served to demonstrate that several factors seem to influence laryngeal airway resistance, including intensity of phonation, the presence of a voice disorder, age and sex.

Leeper and Graves (1984) found that intensity influences laryngeal airway resistance. They collected laryngeal airway resistance data on 15 young women aged 23 to 32 years, over a two-day period, with two sampling sessions per day. The overall mean R_{law} for this group of subjects was 38.3 ± 9.24 cm H_2O/LPS. No significant differences in resistance values were found within or between days for the subjects. The values obtained for women in the this study were similar to the mean values obtained for men by Smitheran and Hixon (1981). However, the mean values were slightly lower than the resistance values reported by Langhans (1981; 46.1 cm H_2O/LPS) and by Shaughnessy, Lotz and Netsell (1981; 43.5 cm H_2O/LPS) and higher than those reported by Stathopoulos, Hoit, Hixon, Watson and Solomon, (1991; 34.53 ± 13.62 cm H_2O/L/s). Overall variability of R_{law} in the Leeper and Graves (1984) data was larger than that found by Smitheran and Hixon (1981) but was consistent with that reported by Langhans (1981) and Shaughnessy, Lotz, and Netsell (1981). Phonation data were also collected during a controlled-intensity condition by Leeper and Graves (1984). Subjects were required to monitor the loudness of their phonation on a sound level meter. The range and variance of laryngeal airway resistance decreased significantly when intensity was controlled at 75 dB SPL. Leeper and Graves state that "from the present results it becomes apparent that in considering patients with vocal fold pathologies, which by their very nature will affect resistance values, some control factors must be employed during vocal function analysis" (1984, p.161).

The findings reported by Leeper and Graves (1984) are consistent with those found by Isshiki (1964) in his study of the regulatory mechanism for vocal intensity. Isshiki found that at low frequencies, as intensity increased, airflow remained relatively unchanged or actually decreased while glottal resistance increased. However, at high frequency phonation levels, as vocal intensity increased, airflow rates increased greatly while glottal resistance remained relatively unchanged. The data of Leeper and Graves showed that laryngeal airway resistance was greatest when intensity was high (80 dB) and pitch was within midrange, yet laryngeal airway resistance exhibited little variability when subjects maintained a stable vocal intensity (75 dB). The results of Leeper and Graves' study suggest that intensity is an important variable to be considered during the sampling of data for the computation of laryngeal airway resistance as it affects the stability of the resistance measures obtained.

Age has also been found to influence laryngeal airway resistance. In children, laryngeal airway resistance changes as they grow older. The results of an investigation completed by Lotz and Netsell (1986) indicated that it is not until early adolescence that children develop adult-like control of the vocal tract in speech. Estimates of subglottal air pressure and laryngeal air flow during vowel production obtained from 70 children aged 3 to 12 years and compared with data from 30 adults, indicated that air pressure progressively decreased with age while airflow progressively increased with age.

Two studies were undertaken to determine whether laryngeal airway resistance is influenced by the aging process in adults (Melcon, Hoit & Hixon , 1989; Hoit & Hixon, 1992). Melcon, Hoit and Hixon (1989) conducted an investigation "to determine if laryngeal valving economy, as reflected in measures of laryngeal airway resistance during vowel production, varies across adulthood" (p. 282). Sixty healthy men were studied, ten from each of six age groups - 25, 35, 45, 55, 65, and 75 years (± 2 years). Results indicated that mean values for resistance measures were relatively similar across the 25, 35, 45, 55, and 65 year old groups, ranging between 35.7 and 40.7 cm H_2O/LPS. Such results fail to suggest any consistent overall relationship between age and laryngeal airway resistance over the forty year span of the first five groups (25, 35, 45, 55, 65) sampled cross-sectionally. However the mean laryngeal airway resistance for the 75-year-old group was found to be only 28.6 cm H_2O /LPS. Pressure measures for the 75-year-old group were similar to those of the other five age groups, whereas measures of flow were found to be significantly higher for the 75-year-olds than for the younger age groups. Melcon, Hoit and Hixon (1989) suggest that-age related changes in the mobility of the cricoarytenoid joints, strength and mass of the laryngeal musculature, and compliance of the lamina propria of the vocal ligaments can interfere with the ability to achieve complete approximation of the vocal folds during voice production. They presumed that the relatively high flow (and low laryngeal airway resistance) noted in the 75-year-old group can be attributed to the changes in laryngeal anatomy that occur as part of the aging process. Such findings have important implications for evaluation and management of voice disorders. Melcon, Hoit and Hixon (1989) suggest that "it is essential that age be taken into account when making judgments regarding laryngeal valving economy from measures of laryngeal airway resistance during vowel production" (p.286).

A more recent study by Hoit and Hixon (1992) was undertaken to determine if laryngeal valving economy, as reflected by laryngeal airway resistance, differs with age in women. Seventy healthy women were studied, ten each at 25, 35, 45, 55, 65, 75, and 85 years of age (± 2 years). Estimates of laryngeal airway resistance were obtained using the method developed by Smitheran and Hixon (1981). A group of 85-year-old women were included in this investigation, as it was believed that changes in laryngeal anatomy occur later in women than in men (Kahane & Beckford, 1991). Mean resistance values ranged from 38.00 to 58.37 cm H_2O /LPS with the mean value for the 45-year-old group substantially lower than that for the other groups. With the exception of the 45-year-old group, the women generally had higher average R_{law} than that calculated for a comparable sample of men (Melcon, Hoit, & Hixon, 1989). As women have relatively smaller airways, it is to be expected that they would have higher resistance to air flowing through their vocal tracts (Hoit & Hixon, 1992). The 25- and 35-year-old women had laryngeal airway resistance values that were on average 15 cm H_2O/LPS higher than those for men of comparable ages; the 55- and 65- year-old women had laryngeal airway resistance values that averaged about 20 cm H_2O /LPS higher than those of men of comparable ages, and the 75-year-old women had laryngeal airway resistance values that were about 25 cm H_2O /LPS higher than those for the 75-year-old men. Laryngeal airway resistance values for the 45-year-old women were similar to those for all but the oldest men whose values these women exceeded by about 10 cm H_2O /LPS. Hoit and Hixon speculated that the "dip" in laryngeal airway resistance in the 45-year-old women might be related to hormonal changes preceding menopause. While Melcon, Hoit and Hixon (1989) had noted lower average laryngeal airway resistance values in the oldest group of men in their study (75-year-old group) in comparison to younger men aged 25-65 years of age, Hoit and Hixon (1992) found that in the women they studied, mean results from the oldest group of women (the 85-year-old group) did not vary significantly from the younger groups of women. However, this female age group had the largest dispersion in R_{law} values. The results of this investigation suggest that it is important for investigators to be aware that measurement of R_{law} will generally be higher in women than in men. Also it is noteworthy that the variance among individual older women can be large, with laryngeal airway resistance values encompassing a wide range (Hoit & Hixon, 1992).

An investigation by Stathopoulos, Hoit, Hixon, Watson, and Solomon (1991) examined respiratory and laryngeal function during whispering in 10 healthy adults. Procedures for laryngeal airway resistance were modelled after those of Smitheran and Hixon (1981). Laryngeal airway resistance measures indicated that whispering involved lower tracheal pressures, higher translaryngeal flows, and lower laryngeal airway resistances than comparable values obtained when speakers used normal conversational voice and intensity. The male subjects' laryngeal airway resistance values averaged 9.63 ± 5.73 cm H_2O/L/s during whispering and 44.55 ± 13.97 cm H_2O/L/s during normally voiced speaking. Female subjects' laryngeal airway resistance values averaged 8.00 ± 1.93 cm H_2O/L/s during whispering and 34.53 ± 13.62 cm H_2O/L/s during speaking.

Netsell, Lotz, and Shaughnessy (1984) used the theoretical basis of the Smitheran and Hixon (1981) method to examine the influence of voice disorders on laryngeal airway resistance. These researchers compared measures of estimated subglottal air pressure and laryngeal air flow during phonation for adults with normal laryngeal function with those for adults with laryngeal abnormalities (e.g., vocal nodules, ulcers, severed laryngeal nerves, and sequelae of head trauma) and found that the presence of a voice disorder affected laryngeal airway resistance. Tasks and measurements used differed only slightly from those recommended by Smitheran and Hixon (1981). Subjects were required to produce three prolongations of the vowel /a/ for approximately two to three seconds and also to produce three series of repetitions of the syllable /pa/, with each series containing approximately seven syllables on one expiratory breath. Netsell, Lotz, and Shaughnessy (1984) found that measures of air flow alone (sampled during the sustained /a/) "offer incomplete and often erroneous impressions of laryngeal aerodynamics" as "laryngeal airflow depends on the magnitude of sub-glottal air pressure as well as on laryngeal airway resistance" (p.400). The investigators found that although all subjects with voice disorders fell outside the normal limits on either the pressure dimension, the flow dimension or both, particular laryngeal abnormalities do not appear to be associated with particular pressure-flow characteristics. Netsell, Lotz, and Shaughnessy (1984) suggest that "an individual's use of the larynx during speech often reflects not only the effects of a laryngeal disorder, but the patient's attempts to compensate for the lesion as well" (p.401).

Netsell, Lotz and Shaughnessy (1984) concluded that "laryngeal aerodynamics provide unique and useful information for the evaluation and treatment of voice disorders" (p.402). The investigators cautioned however that aerodynamic measures are only interpretable in the context of multiple levels of observation. Thus their voice and laryngeal exam included a detailed case history, perceptual ratings of voice, measures of laryngeal aerodynamics during voicing and videofiberoptic laryngoscopy.

Summary of Laryngeal Airway Resistance.

Laryngeal airway resistance, or the extent to which the larynx provides opposition to the airflow through it, provides knowledge about the coupling the larynx permits between the trachea and the pharynx (Smitheran & Hixon, 1981). Laryngeal airway resistance cannot be measured directly by non-invasive means but can be calculated from the ratio of translaryngeal pressures to translaryngeal flows sampled from an utterance that includes alternating production of voiceless plosives and front vowels. Smitheran and

Hixon (1981) developed a non-invasive method for the estimation of laryngeal airway resistance that made clinical assessment of laryngeal aerodynamics a feasible reality.

The Smitheran and Hixon procedure, with minor adaptations, has served as a model for several subsequent investigations of laryngeal airway resistance, and the method's potential flaws have been addressed in the literature (Smitheran & Hixon, 1982; Rothenberg 1982); To date, general investigations of laryngeal airway resistance for several subject groups have resulted in the compilation of a modest data base for clients of both sexes across a wide age range (Hoit & Hixon, 1992; Leeper & Graves, 1984; Lotz & Netsell, 1986; Melcon, Hoit & Hixon, 1989; Netsell, Lotz & Shaughnessy, 1984; and Smitheran & Hixon, 1981). On the basis of these reports, it has become apparent that vocal intensity, age, sex, and the presence of a voice disorder influence measures of laryngeal airway resistance. Extensive clinical applications of the method in treatment studies have not been undertaken, however, though several reports have illustrated the potential for LAR to complement other components of assessment and management plans for disordered voices (Smitheran & Hixon, 1981; Netsell, Lotz & Shaughnessy, 1984). Indeed it would appear that laryngeal resistance measurements could serve as logical complements to perceptual judgements of voice before, during and after therapy for dysphonia.

Perceptual Measures of Voice Quality

Perceptual judgements play an important role not only in describing the perceptual characteristics of a voice disorder at the time of assessment, but also in monitoring changes that may occur in voice quality with treatment. During treatment clinicians make perceptual judgements as to whether a patient's production of a target vocal behavior is a "good" one or not, whether a particular technique is perceived to have had an effect on the quality of a patient's voice, and whether improvement is noticeable from one session to another. The specific methods and standards clinicians use in judging perceptual features of the voice vary extensively, however, and clinicians often do not agree with one another when describing or rating a person's vocal quality. The problems noted in judging perceptual features of the voice and solutions for the disagreements between listeners in judging voice quality deserve mention at this time, because when perceptual measures are used in a reliable manner that yields high between-listener agreement, these measures may be able to complement aerodynamic measures in the assessment of voice.

"The issue of precisely how and why listeners differ in their judgments of vocal quality has received little attention from researchers, despite the potential theoretical and clinical importance such differences may have" (Kreiman, Gerratt, Precoda & Berke, 1992, p.512). In the past it had been assumed that different listeners' ratings of vocal quality should be similar because it was thought that all listeners should share an underlying perceptual set of vocal quality features (Kreiman et al., 1992). Contrary to that assumption, Bassich and Ludlow (1986) reviewed the literature on this topic and found that levels of interrater reliability varied substantially across perceptual measures utilizing scales, listener groups and voice samples. For example, on a 5-point equal-appearing interval scale, Deal and Emanuel (1978) found 80% of ratings within ± 1 scale value when 11 graduate students rated both normal and simulated "rough" vowels. In an investigation by Yumoto, Sasaki, and Okamura (1984), eight laryngologists rated the hoarseness of 87 voices using a 4-point equal-appearing interval scale. Correlations between raters ranged from 0.51 to 0.79. Bassich and Ludlow (1986) themselves reported a mean intraclass correlation of 0.71 with a range of 0.19 to 0.96 when 4 inexperienced but well-trained listeners rated dysphonic voices using a 13-point scale. Based on a review of interrater reliability and agreement which yielded varying levels, these investigators argued that listeners require substantial clinical experience in order for their ratings to agree.

Kreiman, Gerratt and Precoda (1990a) used a multidimensional scaling technique to examine the effects of listener groups and speaker populations on perceptual strategies. Both experienced clinicians and naive listeners attended to different cues when judging the same set of voices. Furthermore, individual clinicians appeared to differ greatly in the attention they gave to various acoustic parameters. A study by Kempster, Kistler and Hilldenbrand (1991) examined relative subject weights for two three-dimensional scaling solutions. The task was completed by 15 clinical trainees who needed to match sets of dysphonic voices. Good agreement was found on one dimension, while listeners did not agree on the importance for the other two dimensions. Kreiman et al. (1992) analyzed the perceptual strategies used by individual listeners. Sixteen listeners (10 expert, 6 naive) judged the dissimilarity of pairs of voices drawn from both pathological and normal populations. Correlations between individual listeners' dissimilarity ratings were low. However "scaling solutions indicated that each subject judged the voices in a meaningful, reliable way, and listeners differed more from one another in their judgments of the pathological voices than they did for normal voices" (p.512). Kreiman et al. (1992) argue that in general, "individual listeners deviate from an average perceptual strategy, at least with respect to the relative importance granted to different vocal features" (p.513). However these investigators suggest that "it is unclear whether listeners in the studies attended to the same things when judging the quality of a given voice or if a given listener consistently relied on the same characteristics when judging the quality of different voices" (p.513). They concluded that perhaps "if listeners differ in perceptual strategies but they each judge voices in ways that are internally consistent and reasonable, then traditional standards for interrater reliability must be rethought" (p.512).

An extensive literature review by Kreiman, Gerratt, Kempster, Erman and Berke (1993) of many perceptual studies using a variety of perceptual tools indicated that both intrarater and interrater reliability fluctuated greatly from study to study. These researchers' own data found that ratings of "vocal roughness" varied widely across clinicians, with a single voice receiving almost all possible ratings. Kreiman et al. (1993) also argued that no model or theoretical framework currently existed to explain the variations. They proposed several possible influences which may have contributed to the variability in ratings including listeners' backgrounds and biases, the task used to gather ratings, interactions between listeners and tasks, and random error.

Kreiman et al. (1992) suggested that it is possible that expert listeners actually store examples of typical perceptual qualities (e.g., rough voices). These authors express the need for using explicit comparisons among voices in gathering perceptual judgments. Many commonly used perceptual tasks (e.g., direct magnitude scaling & voice ratings using equal-interval scales) involve comparing voices to stored representations (i.e., internal standards or memory traces). "These standards may fluctuate from trial to trial, and in any event are not directly available to the listener or the investigator" (Kreiman et al., 1992, p.519). Consequently "calibration of a listener's behavior is not possible" (p.519).

In consideration of the difficulties presented when listeners are asked to compare voices to their "stored" representations in tasks involving direct magnitude estimation or the use of equal-interval scales, it appears that the presentation of paired comparisons of voice samples may be the safest method of obtaining reliable perceptions of voice quality because it provides a standard for listeners. In paired-comparison tasks, a listener is presented with two auditory stimuli. The listener must compare the second stimulus to the first one (Silverman, 1993). In this way, the stimulus presented first functions as an "anchor" for listeners. Listeners do not need to compare voices to an internal standard or a memory trace but can compare the second voice sample in a stimulus pair to the "anchor" stimulus presented first. The choice of using paired comparisons also seems to be a logical one relative to procedures used in a clinical setting. Clinicians frequently listen to successive recordings of voice samples from the same person, or before-after treatment samples and judge whether the voice quality in one sample is better or worse than that in the other.

Summary of Perceptual Studies.

Perceptual judgements of voice quality by clinicians tend to be highly subjective. In the past it had been assumed that different listeners' ratings of vocal quality should be similar because it was thought that listeners should share an underlying perceptual set of vocal quality features (Kreiman et al.,1992). Data contrary to that assumption were reported by Bassich and Ludlow (1986) and Kreiman et al. (1993), who found that levels of interrater reliability varied substantially across various perceptual measures. Furthermore, listeners tend to differ more from one another in their judgements of dysphonic voices than they do for judgements of normal voices (Kreiman et al., 1992 ; Kreiman et al.,1993), and they require substantial clinical experience in order for their ratings to agree (Bassich & Ludlow, 1986). Kreiman et al. (1993) suggest that the large variability present among listeners' ratings may be caused by listeners' backgrounds and biases, the task used to gather ratings, interactions between listeners and the tasks, and random errors. Many commonly used perceptual rating methods (e.g., direct magnitude scaling & equal interval scaling) involve comparing voices to stored representations such as internal standards or memory traces. In these tasks perceptual judgements may lack consistency because listeners are required to rely on echoic memory and perceptual biases, which may be unreliable, in order to compare voices against an internal standard. In consideration of the potential for poor reliability across judgements that rely on memory or perceptual sets, pair-wise comparisons appear to be a preferable choice for perceptual ratings that involve multiple listeners and samples of vocal quality.

Relationships between Aerodynamic and Perceptual Measures

To date, the relationship between aerodynamic and perceptual of vocal function has not been explored in great detail. Studies which have involved comparisons between ratings of vocal quality with aerodynamic measures have been limited.

Harris (1971) used an 8-point severity scale, namely the Wilson Voice Profile (Wilson,1972), to rate the voices of 24 children with vocal nodules and compared the severity of -2+2 (air loss and laryngeal tension appearing simultaneously) laryngeal function ratings with subjects' mean airflow rates during production of vowels. Results indicated a product-moment correlation coefficient of 0.56.

Moran and Gilbert (1984) compared listener ratings of speech samples obtained from a group of normal individuals and a sample of individuals with voice disorders using the laryngeal function (tension and air loss) and pitch categories on the Wilson Voice Profile with several aerodynamic and acoustic parameters. Aerodynamic measures included rate of airflow and volume flow per syllable. In connected speech samples, tension ratings were found to be significantly positively correlated with volume flow per syllable (r = 0.47) and "for sustained vowel and syllable productions, laryngeal tension ratings were significantly positively correlated with volume flow per syllable" (r = 0.59) (p.253). Air loss ratings were significantly positively correlated with volume flow per syllable in connected speech samples (r = 0.45) and in sustained vowel and syllable productions (r = 0.62). For sustained vowel and syllable productions, air loss ratings were also significantly positively correlated with rate of airflow (0.45). In addition, severity ratings of -2+2 were significantly positively correlated with volume flow per syllable (0.58) in connected speech samples and with rate of airflow (0.44) and volume flow per syllable (0.72) in sustained vowel and syllable productions.

The results of these investigations by Harris (1971) and Moran and Gilbert (1984) suggest that a relationship exists between aerodynamic measures and ratings of voice quality. However, the correlations between aerodynamic measures and perceptual ratings reported exhibited a variety of values, ranging from a weak positive correlation of 0.44 to a fairly strong positive correlation of 0.72 with an overall mean correlation of 0.54. Limitations of the perceptual measures in these studies included the use of a voice rating tool which required listeners to be trained in using it, and required internalized representations of various degrees of voice quality disorders based on experience with the rating tool. Limitations in the area of aerodynamics included the fact that only air-flow was measured. Concomitant subglottal pressure and the overall resistance to the airflow travelling through the larynx were not measured.

A single study specifically investigated the relationship between the aerodynamic measures of subglottal pressure and airflow with perceptual ratings of voice quality in patients with voice disorders. Netsell, Lotz, and Shaughnessy (1984) compared measures of estimated subglottal air pressure and laryngeal air flow during phonation for adults with normal laryngeal function with those for adults with laryngeal abnormalities and found that the presence of a voice disorder affected laryngeal airway resistance. All subjects with voice disorders fell outside the normal limits on either the pressure dimension, the flow dimension, or both. Voice quality ratings in the Netsell, Lotz and Shaughnessy project were made by a panel of three experienced speech pathologists using a seven-point equal-interval scale. Listeners were asked to rate three deviant voice quality dimensions (breathy, rough, and strained). In addition, listeners were asked to rate the overall severity of the voice disorder. Although listeners rated voices individually, the criterion score for each perceptual dimension was the mean of the three listeners' scale values. Five speech samples were rated a second time to compare interlistener and intralistener reliabilities. In rerating the five samples, a given listener assigned the same scale point in only 47 per cent of the instances (intralistener reliability). Interlistener rating scores also differed substantially. One extreme example occurred when listener 1 rated one speech sample as a "6" in terms of strained quality, while listener 3 rated it as a "2". "Using the mean of the three listeners' ratings as the criterion measure, the rerated samples generally were within one scale point of each other" (Netsell, Lotz & Shaughnessy, 1984, p.399). The resulting mean difference for the 15 rerated samples was slightly more than half a scale point.

When Netsell, Lotz and Shaughnessy (1984) compared listeners' perceptual ratings with the aerodynamic measures of subglottal air pressure and air flow they concluded that there appeared to be a consistent relationship between aerodynamic measurement and perceptual measurement. Subjects' voices were generally perceived as "breathy" by listeners when pressure was within normal limits and flow exceeded 400 cc/second. Combinations of low flow as well as excessive pressure (pressures being greater than 10 cm H_2O), resulted in the perception of a "strained voice." Combinations of high flow and excessive pressure resulted in the perception of a "rough voice". However, the validity of the correlations between the two measures needs to be questioned because reliability among listeners was so poor on the perceptual rating task.

Summary of Relationships between Aerodynamic and Perceptual Measures.

The few investigators who have studied the relationship between aerodynamic and perceptual measures have argued that there is indeed a relationship between the two. The validity and reliability of these investigations have been limited in several ways however. Past studies have been few, have used rating scales which require listeners to have a "stored representation" of what can be considered normal voice quality, have exhibited poor interlistener reliability, or have not necessarily calculated the overall resistance to the larygeal valving process. In addition the correlation values that reflect the relationship between aerodynamic and perceptual measures have varied considerably from one study to another, and have generally not been very strong. Finally a major limitation of past investigations that have examined the relationship between aerodynamic and perceptual measures is the fact that the correlation has not been explored clinically.

Thus an investigation which examines both aerodynamic and perceptual measures in clinical trials is warranted.

Conclusion of Literature Review

Patients with vocal nodules appeared to be potentially useful subjects for an investigation of changes in laryngeal airway resistance as a result of voice treatment, due to the high incidence of vocal nodules and the fact that voice therapy is recommended as the primary method of reducing the underlying hyperfunctional vocal behaviors (Boone & McFarlane, 1988). Because structural and functional changes are presumed to occur gradually during the course of treatment, (i.e., vocal nodules are resorbed or diminish; more complete vocal fold approximation is achieved), and patients are learning new ways to use their vocal mechanisms, it seems logical that both perceptual and aerodynamic measures might reflect changes in laryngeal function.

The development of a noninvasive method for estimating laryngeal airway resistance by Smitheran and Hixon (1981) provided a technique that can be used to monitor treatment effects at the aeromechanical level of voice production, although few well-controlled treatment studies have been reported that have used laryngeal airway resistance as a dependent variable.

In voice assessment and treatment clinicians are continually involved in making perceptual judgements about a patient's voice quality. Unfortunately perceptual judgements of voice quality by clinicians tend to be highly subjective. Various experiments involving perceptual judgements have indicated that levels of interrater reliability varied substantially across various perceptual measures. As well, many commonly used perceptual rating methods (e.g., direct magnitude scaling & equal-interval scaling) involve comparing voices to stored representations such as internal standards or memory traces which may be unreliable. Paired-comparison listening paradigms appear to be more reliable.

A few investigations have studied the relationship between aerodynamic and perceptual measures as indicators of laryngeal function. Although some of these studies have suggested that aerodynamic measures and perceptual ratings can be correlated, the reliability and validity of these suggestions remain to be adequately tested in clinical trials.

Statement of Problem.

A review of the literature indicates that estimates of laryngeal airway resistance in individuals with disordered voices differ from those of subjects with normal voices. Although reports in the literature have suggested that this measure can be useful for evaluation and management of dysphonias, few examples of such a practise have been reported. If laryngeal airway resistance changes as a result of treatment, such information might provide clinicians with a useful objective measure of treatment efficacy to accompany perceptual judgements of change in voice quality, but unambiguous evidence of the salience of LAR as an indicator of a treatment effect or as a correlate of voice quality has yet to be demonstrated.

Statement of Purpose.

The purpose of this study was to answer the following research questions:

(1) Is there a change in estimates of laryngeal airway resistance obtained from individuals with dysphonia due to vocal nodules across an experimental period of voice therapy?

(2) Is there a relationship between estimates of laryngeal airway resistance and perceptual judgements of voice samples obtained from individuals with vocal fold nodules across an experimental period of voice therapy, utilizing aerodynamic and perceptual measures obtained weekly from the last week of the baseline phase through the last (eighth) week of the treatment phase?

CHAPTER III. METHODOLOGY

Design.

This research took the form of a single-subject A-B and correlation design, implemented on three subjects with vocal nodules. The A-B phases of the study allowed for adequate pre-treatment tracking of the dependent variables (estimates of laryngeal airway resistance and samples of voice quality) during a baseline period (phase A), and continuous documentation of their behavior across a period of treatment (phase B).

Utilization of a single-subject design enabled the investigator to track the experimental targets from one treatment session to another (McReynolds & Kearns, 1983). Data sampled in the pre-treatment (baseline) phase served as a control for the same behaviors sampled in the treatment phase. By virtue of the single-subject design, each individual was compared only to herself; the investigator needed to contend with within-subject variance, but intersubject variability was not an experimental problem.

Experimental Variables.

There were two dependent variables: estimates of laryngeal airway resistance from aeromechanical recordings sampled throughout the experimental period, and perceptual ratings of voice quality on digital acoustical recordings obtained from nine sample periods, one from the last week of baseline and from each week of the eight-week treatment period. A repetitive utterance train utilizing the syllable /pi/ was used to sample the aeromechanical and acoustical data simultaneously. The independent variable was voice therapy.

Experimental Phases.

The A-B components of the research design constituted two experimental phases: baseline and treatment. The baseline phase consisted of a total of 12 sessions over a four- week period, while the treatment phase consisted of 24 sessions over an 8 1/2 week period. Because the experiment continued into the summer months, each subject's experimental period was interrupted by a period of vacation approximately a week in length. Subject #1's vacation ocurred between her sixth and seventh week of treatment. Subject #2's vacation ocurred between third and fourth week of the treatment phase, and subject #3's vacation ocurred between the last week of baseline and the first week of treatment.

(1) Baseline

During the four-week baseline period, voice recordings for perceptual judgements and pressure/flow recordings for estimates of laryngeal airway resistance were obtained simultaneously three times per week. No treatment was provided during this time.

(2) Treatment

During the 8 1/2-week treatment phase, voice recordings for perceptual judgements and pressure/flow recordings for estimates of laryngeal airway resistance continued to be obtained simultaneously at the beginning of each treatment session approximately three times per week. During the remainder of each session in this phase, voice therapy was provided to the subject by the investigator in the role of clinician.

The experimental design was organized into phases divisible by four weeks for several reasons. It was expected that a four-week pre-treatment phase would provide an opportunity to establish a stable baseline and to observe the effects of factors which might possibly influence voice quality or laryngeal airway resistance such as the menstrual cycle, colds and practice effects. It was also expected that an eight-week treatment phase would allow sufficient time for subjects' voices to respond to therapy (Colton & Casper, 1990) and would provide further opportunity to monitor menstrual cycle effects, if any, on the dependent variables (Ward, 1992).

Subjects.

Potential subjects were patients diagnosed with vocal nodules in the Voice Clinic at the Glenrose Rehabilitation Hospital. The presence of vocal nodules was documented via a flexible fiberoptic video nasendoscopic (FFVN) laryngeal examination performed by an otolaryngologist. The research study was offered to these patients as an alternative to waiting for treatment at the Glenrose. Interested participants were asked by clinic personnel to sign a form allowing the primary investigator to contact them by telephone and provide further information regarding the study (Appendix A). In addition interested potential subjects received an information package describing procedures of the research study (Appendix B).

Eligible subjects must have abstained from smoking for the last two years, must not have been suffering from acute or chronic respiratory illness, and must have had no history of neurological disease, chronic allergies, chronic indigestion, or speech and hearing disorders. Subjects must also have abstained from taking anti-psychotic drugs, thyroid replacement hormones, androgenic agents, anti-hypertensive agents, diuretics, beta blockers and antihistamines for two weeks prior to participating in the study. All of these factors have known or suspected effects on the voice and may negatively influence the success of voice therapy (Sataloff, Lawrence, Hawkshaw, & Rosen, 1991). In addition subjects who were professional singers or actors were not eligible to participate in the study because of the possibility that these individuals may be able to change their laryngeal airway resistance through compensatory strategies even when a voice disorder is present. Although one potential subject reported chronic allergies to trees, dust, grass and some foods, it was decided that she would be eligible to participate in this study. The subject reported that the allergies caused her to have somewhat "watery" eyes at times, but nasal congestion, excessive mucous and post-nasal drip were not usually problematic. In addition she strongly felt that her allergies were not the cause of her voice problem, as the allergies were long-standing while the duration of her voice problem was only 1 1/2 years. Her allergies were managed with shots which she received approximately every three weeks and with antihistamines whenever allergic symptoms flared. The subject felt that she could manage without antihistamines, and she was accepted into the experiment with the understanding that she would abstain from taking this medication during her participation in the research. In addition the researcher allowed for a two-week lapse between the time this subject ceased to take the antihistamines and the time her baseline phase began in order to ensure that the medication was no longer active in her body.

All potential subjects were also required to pass a pure-tone hearing screening at 35 dBHL for the frequencies 500, 1000, and 2000 Hz (ASHA, 1978) administered by the investigator utilizing a portable audiometer (Maico-MA-41; calibrated to ANSI-1969 standards). A hearing screening was a necessary procedure, because severe hearing loss over a long period of time can result in a change in habitual vocal pitch levels and pitch variability (McGlone & Hollien, 1963). A hearing loss may also interfere with an individual's ability to monitor his or her voice. Thus hearing loss may influence success in voice therapy (Boone & McFarlane, 1988).

Four females between the ages of 15 and 38 years of age were accepted to participate in the study. All subjects were fluent speakers of English. One of the four subjects completed the baseline phase of the experiment but was excused after the third week of treatment due to inconsistent attendance, and her name was placed back on the treatment waiting list at the Glenrose Rehabilitation Hospital. Three subjects completed the experiment. A summary of information about them is presented in Table 1 of the Results section.

Procedure

Subject selection.

An initial screening of potential subjects was conducted by the investigator over the telephone (Appendix C) (Ward, 1992). Questions were asked regarding the candidate's history of speech and hearing problems, smoking, respiratory illness, neurological disease, allergies and medications. The telephone screening procedure eliminated unnecessary trips to the research site for further screening procedures by candidates who might be ineligible. Potential subjects were not included if they reported a) a history of speech and hearing problems, b) a history of smoking in the last two years, c) an acute or chronic respiratory illness such as pneumonia or asthma, d) a history of chronic allergies causing nasal congestion, coughing, excessive mucous and post nasal drip, e) a history of chronic indigestion, f) dependence on drugs such as anti-psychotics, thyroid replacement hormones, androgenic agents, anti-hypertensive agents, diuretics, beta blockers and antihistamines, g) a history of neurological disease, and h) being professional singers or actors. Potential candidates who were on medications listed under (f) but were able and willing to discontinue their medication for the duration of the research study were accepted on the condition that a two-week period would elapse between the time they discontinued use of the medication and began the baseline phase.

Persons who did not meet inclusion criteria were informed that they were not eligible to participate in the research study, and the reasons for ineligibility were explained to them. They were thanked for their time and interest and were assured that they would remain on the treatment waiting list at the Glenrose Rehabilitation Hospital for voice therapy. The Voice Clinic at the Glenrose Hospital was notified that the person was ineligible to participate in the study in order to ensure maintenance of her position on the hospital treatment waiting list.

Subjects who did meet inclusion criteria during the telephone screening procedure were informed that they were eligible to participate in the research study and an initial appointment was scheduled to begin the experimental protocol.

Experimental Protocol.

At the beginning of the first baseline session subjects who were deemed eligible on the basis of the telephone interview were re-apprised of the purpose and procedures of the study, and their written consent was obtained (Appendix D). A voice case history was taken to reconfirm that each individual met all the selection criteria (Appendix E), and each subject's hearing was screened. During the four-week baseline phase subjects visited the experimental site three times per week for sessions lasting approximately 5-15 minutes. During each of these sessions aeromechanical data and audio recordings were obtained, and then the subject was excused.

Following the four-week baseline phase, the treatment phase commenced. During the treatment phase subjects continued to visit the experimental site approximately three times per week, and aeromechanical data and audio recordings continued to be obtained at the beginning of each session. However during this phase of the experiment voice therapy was provided at the conclusion of data collection. Hence, sessions during the treatment phase lasted approximately 50-60 minutes. Detailed procedures for voice therapy are described in Appendix F.

Because the menstrual cycle has been reported to have effects on the pitch and pitch variability of the voice (Higgins & Saxman, 1989), subjects had been asked in the initial information packet to report the commencement and duration of their menses to the investigator throughout the experimental period. Subjects were also asked to report when they were suffering from a cold or allergies or when they were taking medication of any kind. These reports enabled the investigator to monitor any effects of medications, the incidence of upper respiratory congestion and drainage associated with colds and allergies and possible effects of the menstrual cycle on vocal function across the baseline and treatment phases.

At the end of the treatment phase, each subject underwent a second flexible fiberoptic videonasendoscopic laryngeal examination performed by the same otolaryngologist at the Glenrose Rehabilitation Hospital, the Charles Camsell Hospital or the physician's office. This provided an opportunity for the experimenter to document any visible structural changes in the subjects' vocal nodules.

During the experimental period for each subject, data acquisition and analysis of the dependent variables occurred in the following manner: (1) Aeromechanical data, in the form of subglottal air pressure (cm H_20) and translaryngeal flow (L/s) were collected at every experimental session, and laryngeal airway resistance was computed by dividing pressure values by flow values. (2) Acoustical data were acquired simutaneously with the aeromechanical measures as analog tape recordings that were subsequently low-pass filtered, digitized and edited for playback to listeners. (3) The digital audio recordings were randomized and presented to experienced listeners in pairwise comparisons for perceptual judgements of voice quality between members of a pair. (4) The aeromechanical data were graphed to allow for visual inspection. (5) Correlations between differences in laryngeal airway resistance and judgements of voice quality were computed for nine points over the experimental phase (one from the fourth week of baseline, and one from each week of the treatment phase). Further details of these procedures are described below.

Laryngeal Airway Resistance Data Collection and Measurement.

Data required to estimate laryngeal airway resistance were obtained via the PERCI-PC system (version 2.0, Microtronics, Inc., Chapel Hill, NC.) operating in MS-

DOS on a dedicated microcomputer (Zenith 386). The PERCI-PC system allows its users to sample airway pressures and flows simultaneously during the production of speech. Calibration of the PERCI-PC was completed according to manufacturer's specifications daily prior to its use with each subject. Care was taken to ensure that instruments were warm and functionally stable prior to the commencement of calibration procedures. Pressure calibration was accomplished against a U-tube water manometer at 10 cm H_2O. Flow calibration at 0.250 L/s was accomplished using a Gilmont rotameter arranged in series with a heated Fleisch-type pneumotachometer. A record of scale factors assigned by the PERCI system software for pressure and flow calibration values was maintained across the entire experimental period to ensure that the equipment functioned reliably from session to session. Following calibration, instruments were left on until all testing procedures had been completed on a particular day.

Laryngeal airway resistance estimates were obtained according to the following procedure: Each subject was seated upright and positioned so that the mouth and the nose could be comfortably placed into a full-face anesthesia mask (SCRAM type, Ohio Medical), mounted on the heated pneumotachometer. The distal end of a pressure-sensing catheter in the mask was positioned in the midline of the oral cavity approximately 1.0 cm behind the central incisors. The subject was required to produce the syllable /pi/ seven times on one breath while connected to the instrumental array. Figure 1A illustrates the placement of the mask and catheter on the subject, while Figure 1B illustrates the complete instrumental array. The anesthesia mask was required to collect airway-opening flow produced during the vowel /i/ and channel it to the pneumotachometer; oral pressure produced during the /p/ was sampled by the pressure-sensing catheter. Care was taken to ensure that the interface between the mask and the subject's face was maintained firmly and without gaps that could leak during the data recording.

Subjects were informed of the nature of the task to be performed and were instructed to model their utterances after those of the investigator. Subjects were asked to produce three syllable trains of the utterance /pi/ at a rate of approximately 1.5 syllables/s as per the method of Smitheran and Hixon (1981). For each utterance train, subjects were advised to take a breath they imagined was deep enough to count outloud from one to seven. This utterance rate was fast enough and the inspiratory depth sufficient enough to eliminate the chance occurrence of intermittent velopharyngeal opening between syllables (Thompson & Hixon, 1978) or separate respiratory gestures (Rothenberg, 1982). Subjects who had difficulty modeling the rate after the investigator were given a chance to practise with the aid of a metronome. In addition subjects were trained to remain within a sound level range of 64-70 dB for the syllable utterance by watching a sound-level meter (Realistic 33-2050) during the data recording. Controlling intensity was a neccessary procedure to ensure that subjects were producing the syllables at a level that was equivalent to normal conversational loudness. Intensity levels may have an effect on resistance values and particularly appear to have an effect on the stability of resistance values that are obtained across recording sessions (Leeper & Graves, 1984).

Three multi-syllable trains of the utterance /pi/ were recorded for each subject at the beginning of each session in both baseline and treatment phases. Subject utterances judged by the investigator to be inappropriate copies of the model for loudness and syllable stress were not accepted, and additional multi-syllable trains of the /pi/ utterance were elicited until three multi-sllyable trains that were appropriate for loudness and syllable stress were obtained. Acoustical recordings of voice during the syllable repetitions that would be used for perceptual judgements were obtained simultaneously with the laryngeal airway resistance samples. (Refer to Acoustical Data Acquisition section).

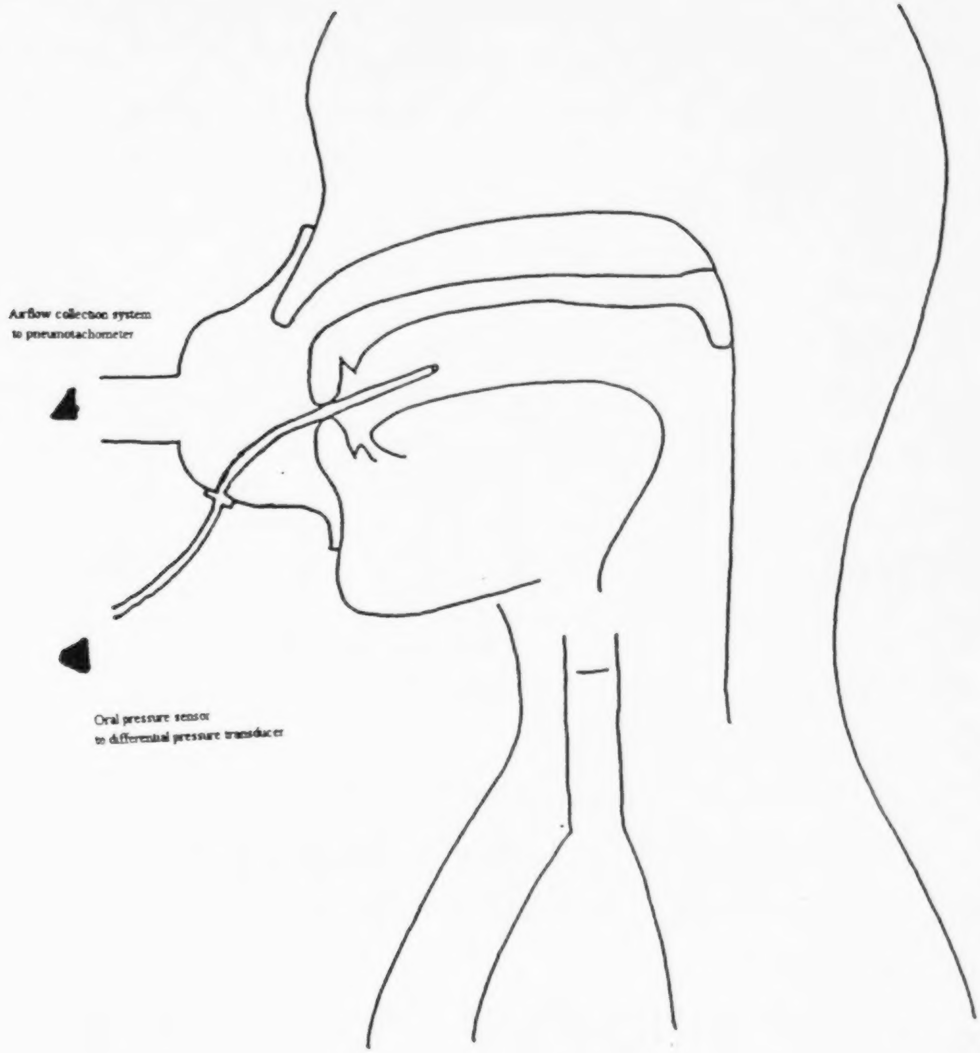

Airflow collection system
to pneumotachometer
Oral pressure sensor
to differential pressure transducer

Figure 1A. Position of Mask and Catheter on the Subject. Reproduced from "PERCI-PC Manual". Copyright 1986 by MicroTronics, Corporation, Chapel Hill, N.C. Reprinted by permission.

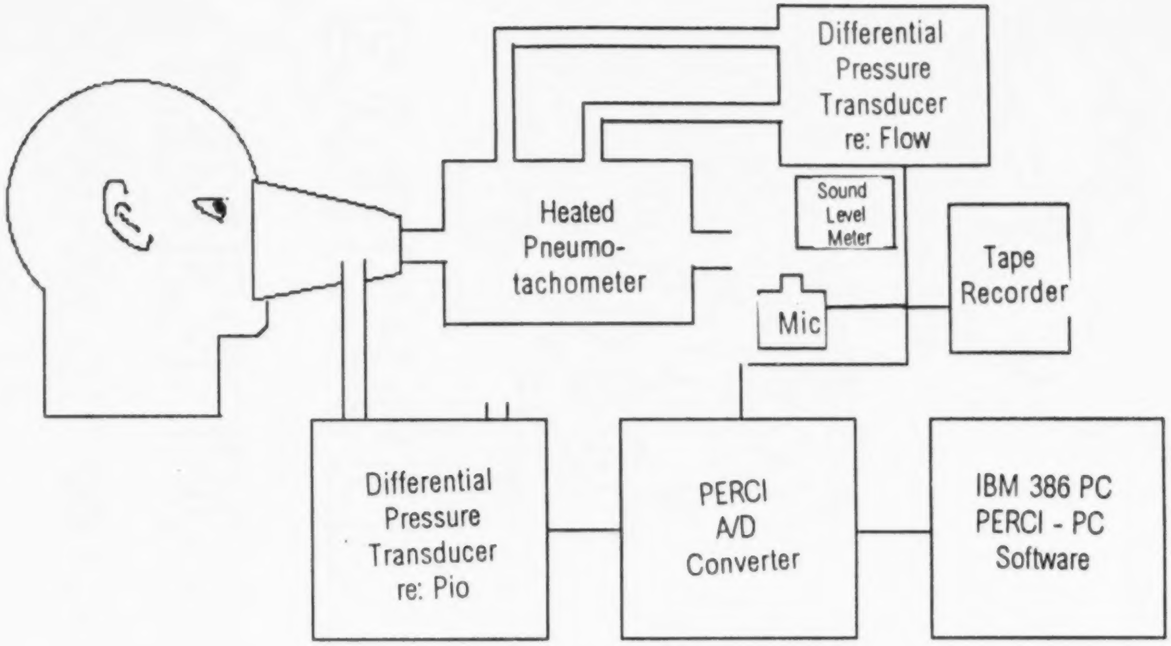

Differential
Pressure
Transducer
re: Flow
Heated
Pneumo-
tachometer
Sound
Level
Meter
Tape
Recorder
Mic
Differential
Pressure
Transducer
re: Pio
PERCI
A/D
Converter
IBM 386 PC
PERCI - PC
Software

Figure 1B. Instrumental Array for Laryngeal Airway Resistance Data Collection. Reproduced from "Laryngeal Airway Resistance in Teachers with Vocal Fatigue". Copyright 1996 by Barbara Kostyk. Reprinted by permission.

The PERCI-PC software (version 2.0), was used to compute laryngeal airway resistance on the utterance samples obtained according to the protocol above. The first and last syllables were not included in measurements in order to eliminate utterance onset and offset effects (Smitheran & Hixon, 1981). A cursor was positioned on the second pressure peak in the utterance series, and three subsequent peaks were automatically selected by the computer software and connected by a horizontal dotted line, as shown in Figure 1C. This line represented an estimate of subglottal air pressure maintained during the portions of the utterance series under scrutiny. After selection of the pressure peaks the cursor automatically appeared on the flow waveform where it was used to mark the portion of the flow trace associated with each vowel in the series, as shown on Figure 1C. When the flow segments corresponding to the subglottal pressure contours were isolated, laryngeal airway resistance was automatically calculated by the computer and displayed at the top of the screen. Because the computer software considered only 3 syllables at a time, each utterance series was measured twice to include five mid-series syllables in the laryngeal airway resistance estimate for a given train. For one subject (# 1), however, only three mid-series syllables were measured as she was unable to produce seven /pi/ syllables on one breath reliably.

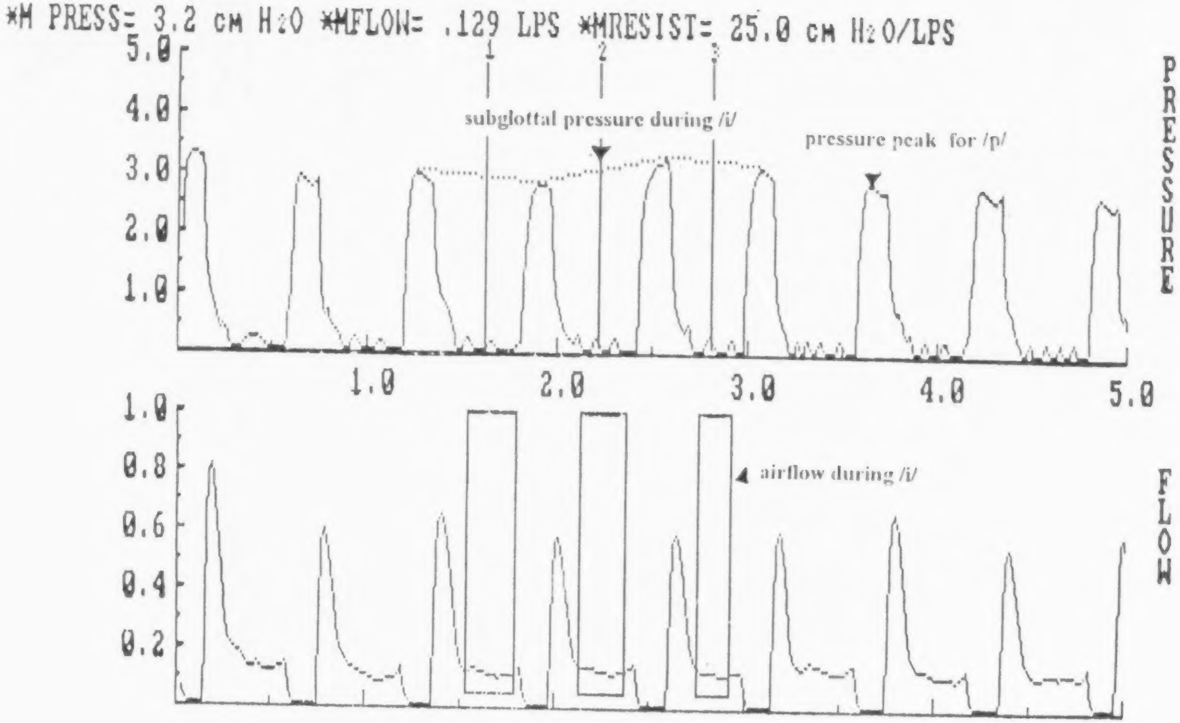


Figure 1C. Laryngeal Airway Resistance Sample Data. Reproduced from "Perci-PC Manual". Copyright 1986 by R. Lutz. Reprinted by permission.

The investigator assessed the reliability of her initial aeromechanical measurements by repeating the measurement procedure for a random sample of 10% of the records obtained for each subject. Laryngeal airway resistance values were calculated from the pressure and flow values (as described earlier) and the difference between the initial values and the repeated measures was calculated and expressed in cm H_2O /L/s. Interjudge reliability was also determined by having a second person familiar with LAR measurements on the PERCI-PC system repeat measurement procedures for the same random sample of 10% of the aeromechanical data. The measurement error was calculated for the differences between the two judges' measurements. Intrajudge remeasurement difference ranged from 0.10 to 6.02 cm H_20/L/s. Interjudge remeasurement difference ranged from 0.01 to 11.1 cm H_20/L/s.

Acoustical Data Acquisition and Analog-to-Digital Conversion.

It was necessary to obtain ratings of subjects' voice qualities in order to compare laryngeal airway resistance data across time with perceptual information about voice quality during that same period. In order to reliably compare the two measures (aeromechanical and perceptual) it was necessary to obtain both sets of measures in the same utterance context. Hence, audio recordings of each subject's productions of the multi-syllable trains were made on a audio cassette recorder (PMD221 Marantz, Chatworth, C.A.), simultaneously with the acquisition of pressure-flow data for laryngeal airway resistance estimation during each session of the baseline and treatment phase. A coding system was used to identify the subject on the tape recordings. In addition audio recordings of each subject reading the first half of the Rainbow Passage were made during each session. During this recording, however, the subject was no longer coupled to the aeromechanical sampling array. Only one sentence of the passage, ("The rainbow is a division of white light into many beautiful colors.") was utilized for perceptual data acquisition.

Taped samples (both the syllable trains and the connected speech samples) from nine moments in the experiment (one from the last week of baseline and one from each of the eight weeks of the treatment phase) were low-pass filtered at 8600Hz (Series 901, Frequency Devices, Haverhill, MA) to eliminate aliasing, and digitized at a 22 kHz sampling rate via the MacRecorder (Farallon Computing, Inc.) at 8-bit resolution on a Macintosh SE-30 computer (Apple Computer, Inc.). The digitized data of the syllable trains were then edited to eliminate onset and offset artifacts just as the corresponding aeromechanical data had been edited. Normally, the first and seventh syllables of each utterance train as well as any extra syllables a subject may have produced, were deleted. For one subject (#1), however, utterance trains that were three syllables in length were utilized. For this subject the first and fourth syllable of each utterance train as well as any extra syllables recorded were deleted. After editing, the data syllable trains were stored as resource files in SoundEdit (version 2.0.1; Farallon Computing, Inc.) for presentation to listeners.

The investigator assessed her reliability for editing the digitized acoustical samples by repeating the editing process for a random sample of two syllable trains from each subject's acoustical data. The measurement error was calculated for the differences between the length of the acoustical sample in milliseconds after the editing process the first time versus the second time. The measurement error ranged from 0 to 2.36 milliseconds for the onset of the acoustical sample and from 0.18 to 24 milliseconds for the offset of the acoustical sample.

Perceptual Data Acquisition.

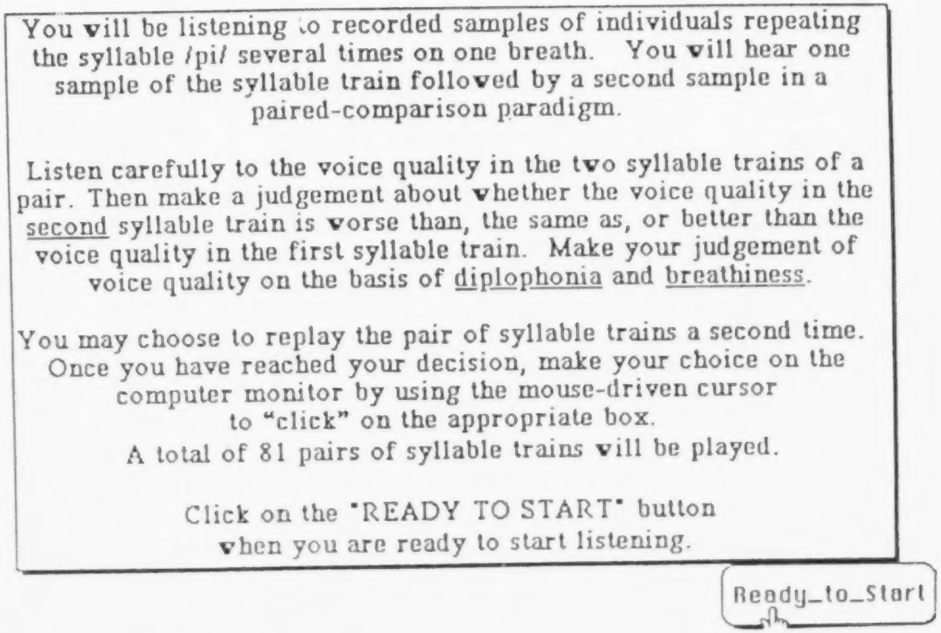
The nine samples of utterance trains and the nine sentences from the connected speech samples digitized from tape recordings and edited as described in the preceding section were randomized by means of a program written in HyperCard (version 2.0, Apple Computers, Inc.) and replayed to listeners in pairwise comparisons from a

Macintosh IIci computer. From the set of nine syllable trains and nine connected speech samples, (one each from the baseline and 8 each from the treatment phase), 81 possible pairings of syllable trains and 81 possible pairings of sentence samples resulted for presentation to listeners. The number of possible pairings was derived by pairing the baseline syllable train with itself and with each of the eight treatment-week samples included as part of the perceptual experiment. Digitized voice samples were utilized from the second syllable train of the third session of each respective week with the exception of two that were contaminated by sound artifacts obtained in the recording process. For these samples (Treatment Wk5 for subject 2, and Treatment Wk 3 for subject 3) the third and the first syllable trains, respectively, of the third session of the week were utilized.

The perceptual judgements of digital samples of each subject's voice were completed in separate listening sessions by three speech-language pathologists experienced in the area of voice disorders. Each listener was seated in front of the computer. The ambient noise level of the room in which the listeners participated was monitored with a sound-level meter (Realistic 33-2050). Listeners rated syllable trains and isolated sentences in separate tasks. Judges listened to each pair of /pi/ syllable trains or pair of sentences once and then had the option of replaying the paired set once more if they required an additional listening opportunity in order to make a judgement. An interval of 1.0 second was presented between the two members of a pair of stimuli. Listeners were asked to judge whether the second member of a pair was better than, the same as or worse than the first member of the pair based on the qualities of breathiness and diplophonia. Listeners made their selection from choices on the computer monitor ("better", "same", or "worse") and responded using a mouse-driven cursor. Following a listener's response an interval of 1.5 seconds elapsed before the presentation of another pair. The order of stimulus pair presentation was randomized by the computer to preclude the possibility of an order-effect on perception of the utterances. Figure 2A illustrates the instructions given to judges on the computer at the outset of the listening task, and Figure 2B illustrates the response screen display that listeners used to make their judgements about each pair of stimuli.

Similar instructions were given for the connected speech samples. In this case the listeners were informed that they would hear voice samples consisting of a familiar sentence presented in paired sets, and they were instructed to rate the sentences in the same manner as the paired syllables using a mouse-driven cursor. Subjects rated 81 pairs of sentences.

Each listener agreed to participate in a second session, in which the entire stimulus pool of syllable trains and sentences was replayed by the computer in re-randomized paired comparisons to assess intrajudge reliability. Percent exact agreements between scores were calculated for the two sets of perceptual ratings completed by each rater for each subject. Interjudge reliability was determined by calculating percent exact agreement among listeners' ratings for the entire stimulus pool of paired comparisons for each subject. Judges required approximately 2 1/2 hours to complete the initial rating session and approximately 2 hours to complete the second rating session.



You will be listening to recorded samples of individuals repeating the syllable /pi/ several times on one breath. You will hear one sample of the syllable train followed by a second sample in a paired-comparison paradigm.

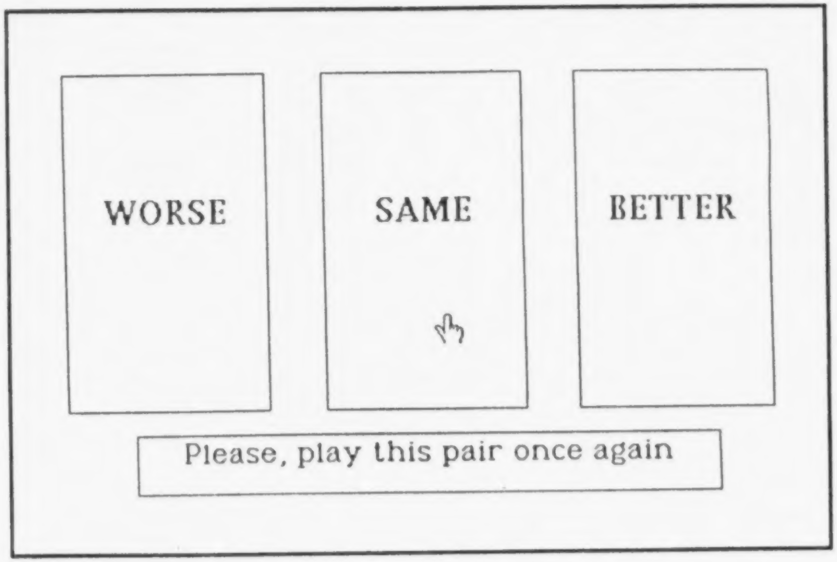
Listen carefully to the voice quality in the two syllable trains of a pair. Then make a judgement about whether the voice quality in the second syllable train is worse than, the same as, or better than the voice quality in the first syllable train. Make your judgement of voice quality on the basis of diplophonia and breathiness.

You may choose to replay the pair of syllable trains a second time. Once you have reached your decision, make your choice on the computer monitor by using the mouse-driven cursor to "click" on the appropriate box.

A total of 81 pairs of syllable trains will be played.

Click on the "READY TO START" button when you are ready to start listening.

Figure 2A. Instructions Presented to Raters.


WORSE
SAME
BETTER

Please, play this pair once again

Figure 2B. Response Screen Display for Raters.

Data Analysis

Data analysis for the aerodynamic measures consisted of visual inspection of the time-series displays of laryngeal airway resistance, subglottal pressure and translaryngeal flow across the baseline and treatment phases for each subject. To assist visual inspection of the data the experimenter constructed a two-standard-deviation band to determine whether there was a clinically significant change in resistance, pressure and flow values across the two phases (Ottenbacher, 1986). To obtain the two standard deviation band, the mean and the standard deviation were computed for the baseline data. A dotted line was drawn to show the point two standard deviations above the mean, and another dotted line was drawn to show the point two standard deviations below the mean. The area between these two dotted lines was termed the two-standard-deviation band, and these lines were extended across the treatment phase. The treatment data series was then scrutinized for signs of two or more data points that exceeded the two-standard-deviation band.

Once the perceptual rating task had been completed by judges, the investigator determined if the order in which a member of a pair was presented to listeners had an effect on the rating of voice quality that it received. T -Tests for paired samples (Bruning & Kintz, 1987) were completed for the ratings of each judge between the two conditions of order (ie., first place in the pair versus second place in the pair) for each subject for both syllable trains and sentences.

Finally, to explore the relationship between the aeromechanical and perceptual data, the experimenter used the Spearman Rank-Order Correlation Coefficient (Seigel, 1956) to correlate the speakers' laryngeal airway resistances with the listeners' ratings of of their voice samples across nine points in the experimental period. The laryngeal airway resistance values were continuous, interval-level data, while the perceptual ratings (1=Worse, 2=Same, 3=Better) were discrete, ordinal-level data. Thus, the Spearman

Rank -Order Correlation Coefficient, the non-parametric equivalent of the Pearson product-moment correlation was chosen to correlate the interval and ordinal level data. Perceptual ratings used for the correlations consisted of the rating values (1=Worse, 2=Same, 3=Better) given to the second member of a pair of syllable trains or sentences in the perceptual experiment. Laryngeal airway resistance values for the correlations were derived from the resistance calculations obtained for the syllable trains: The second laryngeal airway resistance value (which corresponded to the second syllable train of a perceptual pair) was subtracted from the first laryngeal airway resistance value (which corresponded to the first syllable train of a perceptual pair) to yield a single LAR difference value. The perceptual ratings of the sentence samples were correlated with the laryngeal difference values for the syllable trains obtained in the same experimental session as the sentences were recorded. Thus it is acknowledged that the correspondence between laryngeal airway resistance values and perceptual ratings of the syllable trains was a real one (ie., both aerodynamic and perceptual data were obtained for the same stimuli, the /pi/ syllable trains), whereas the correspondence between the laryngeal airway resistance values and the perceptual ratings of the sentences was an artificial one (ie., aerodynamic and perceptual data derived from different stimuli were compared: the resistance data came from the syllable trains, and the perceptual data came from judgements of sentences recorded when subjects were unencumbered by the aeromechanical sampling instruments).

CHAPTER IV. RESULTS

Flexible Fiberoptic Videonasendoscopy

A post-treatment FFVN laryngeal examination completed by the same otolaryngologist who had examined the subjects prior to treatment was used to document any structural changes in the subjects' vocal nodules following voice therapy.

Prior to treatment, subject #1 presented with large bilateral vocal fold nodules, with the nodule on the right side being slightly larger than the nodule on the left side. Actual size of the nodules was not documented at the time of nasendoscopy. At the end of treatment subject #1 showed no more sign of nodules, although her vocal folds appeared somewhat edematous. Prior to treatment, subject #2 presented with broad-based vocal fold nodules that were not well defined. At the close of treatment, subject #2 showed no more sign of nodules. Subject #3's bilateral vocal fold nodules were estimated to be 3-4 mm in size by the otolaryngologist prior to treatment. At the end of treatment, her nodules were judged to be reduced to less than 1mm in size. Table 1 provides a summary of background information on each of the three subjects and provides information about the pre- and post-treatment status of each subject's vocal nodules.

Table 1. Subject Information.

	Subject 1	Subject 2	Subject 3
Age	28	15	38
Occupation	Switchboard operator.	Student	Psychiatric nurse.
General Health	Good	Good	Allergies (dust, grass, trees, and some foods).
Onset of Dysphonia	~ 1.5-2 yrs ago	~1 yr ago	~1- 1.5 yrs ago.
Status of Nodules before Treatment	Bilateral Right nodule larger.	Bilateral Not well defined.	Bilateral ~3-4 mm.
Status of Nodules after Treatment.	Absent. Vocal folds were edematous.	Absent.	Reduced to <1mm.

Laryngeal Airway Resistance

Visual analyses of graphs of laryngeal airway resistance and its components of pressure and flow were used to determine whether resistance changed over the experimental period. Figures 3A, 4A, and 5A show laryngeal airway resistance values in cm H_2O /L/s across time (baseline and treatment) for each of the three subjects. Onset of menses is indicated on the graphs by arrows, and sessions where signs of a cold were present are indicated by small triangles above the session number. Figures 3B, 4B, and 5B show translaryngeal pressure values in cm H_2O across time (baseline and treatment) that correspond to the resistance data for each of the three subjects, while Figures 3C, 4C, and 5C show translaryngeal flow values in L/s across time. Visual analysis of the data across the treatment phase of the experiment inspected the levels of each variable, day-to-day and within-session variability, and overall trends in the data. The average baseline and treatment period values for laryngeal airway resistance, pressure, and flow are indicated on the graphs for each subject. A two-standard-deviation bandwidth based on the baseline data was calculated to aid visual inspection of the treatment data for each subject, and these bandwidths are indicated on the graphs by dotted horizontal lines.

For two subjects, laryngeal airway resistance remained relatively stable across time or changed only moderately. For one subject, laryngeal airway resistance increased remarkably during the treatment phase.

Fig. 3A Subject 1 Resistance

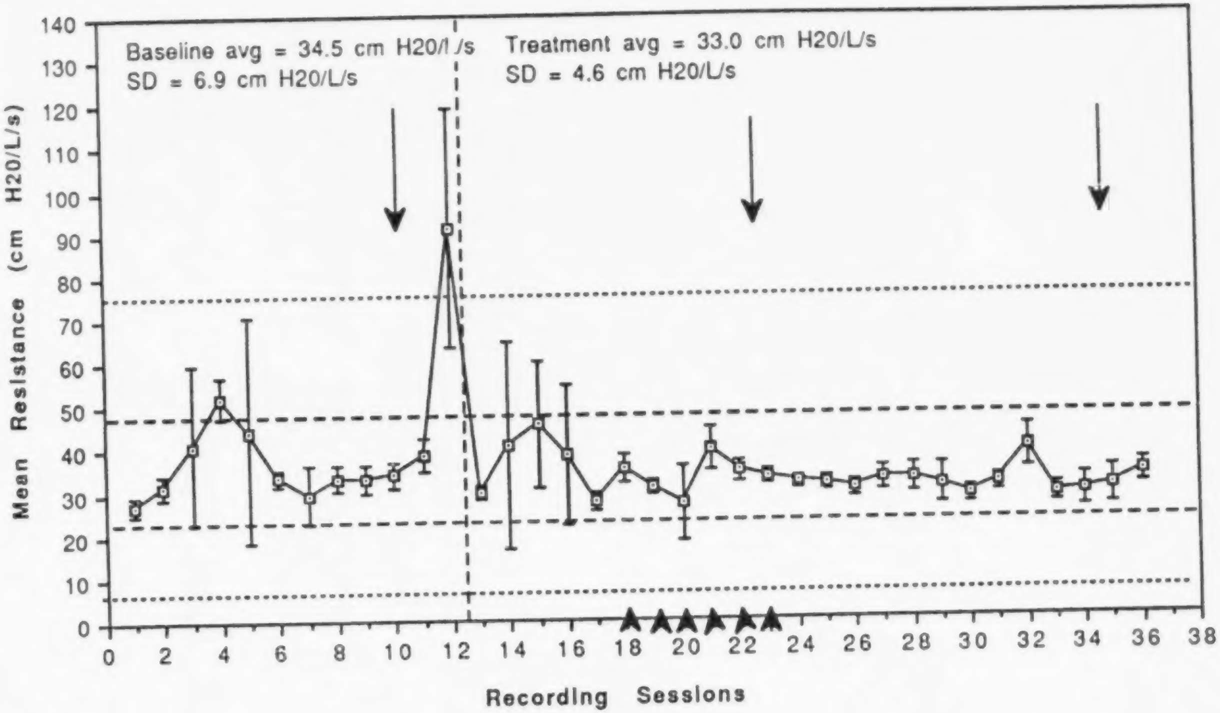


Fig. 3B Subject 1 Pressure

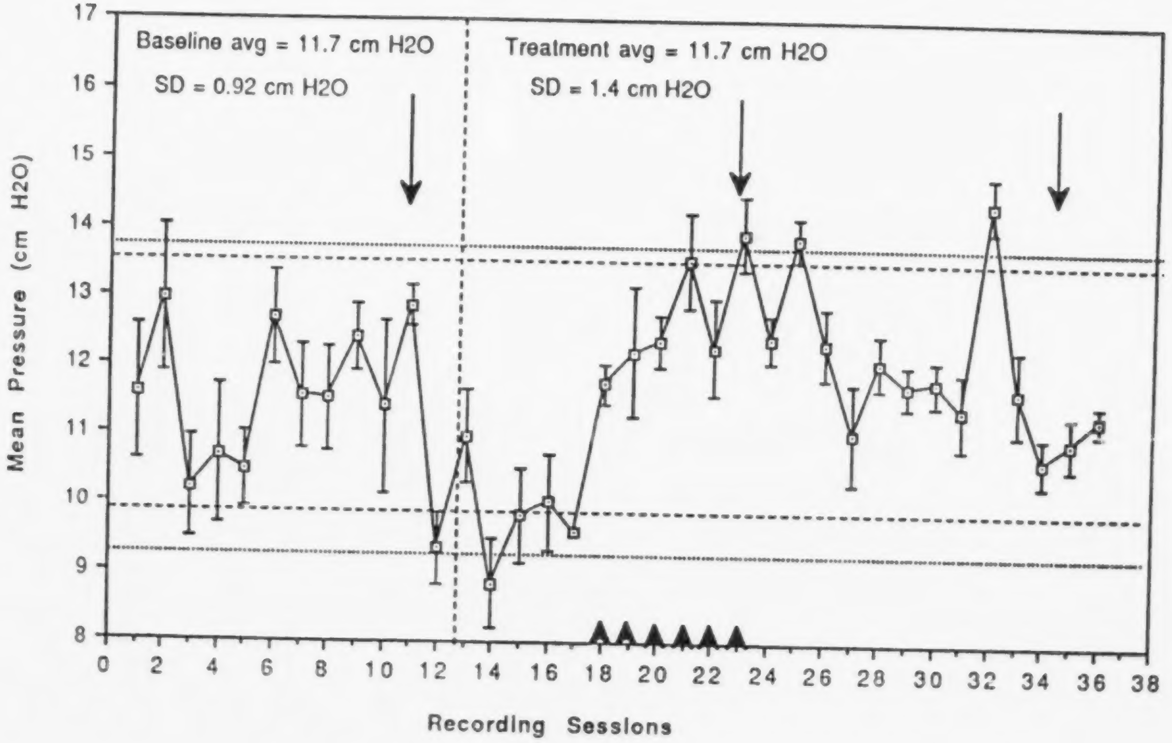


Fig. 3C Subject 1 Flow

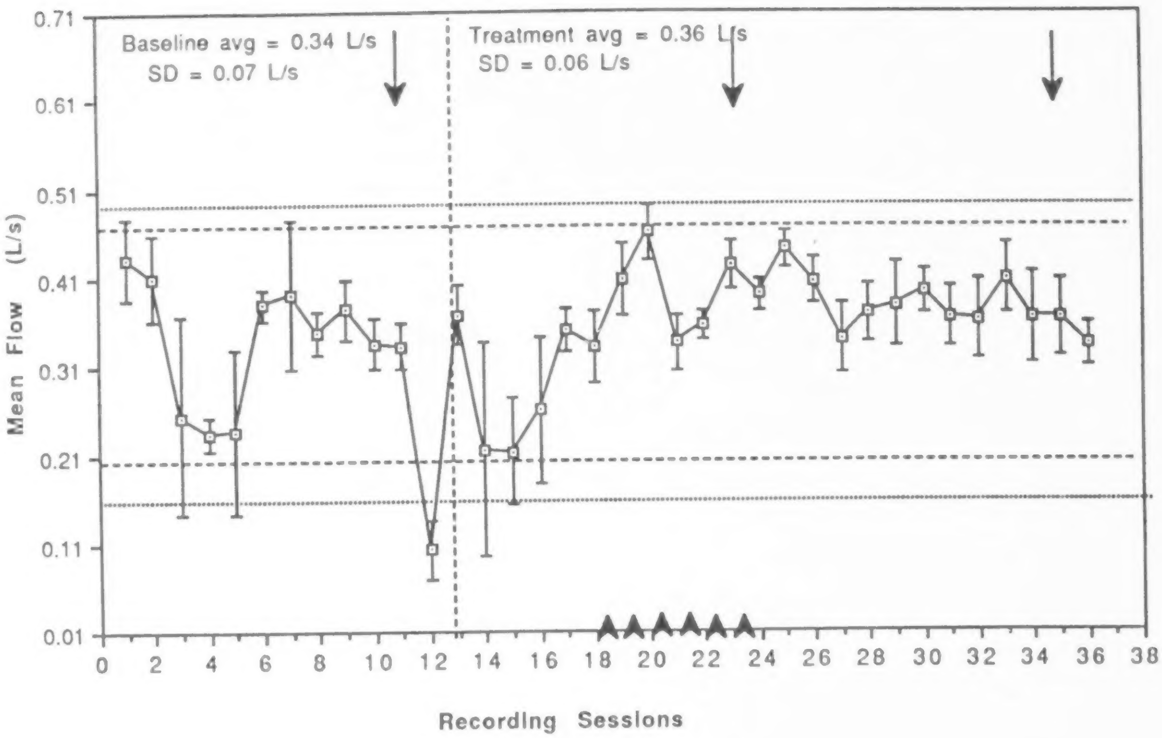


Fig. 4A Subject 2 Resistance

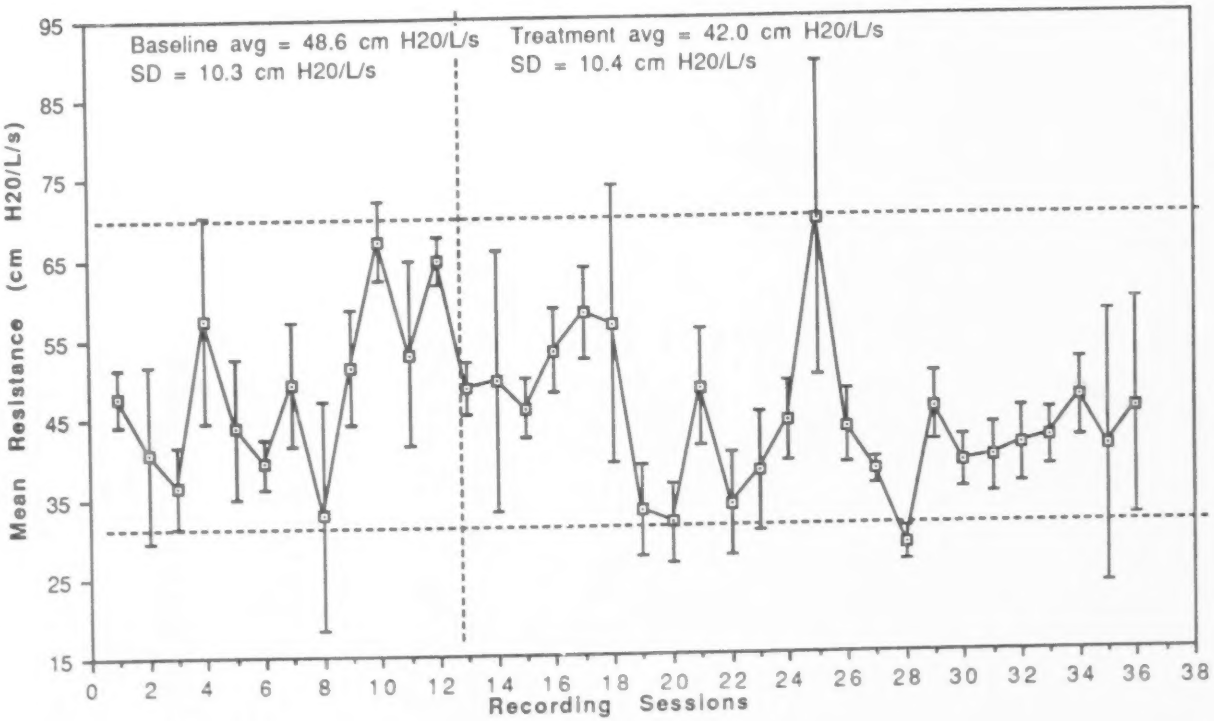


Fig. 4B Subject 2 Pressure

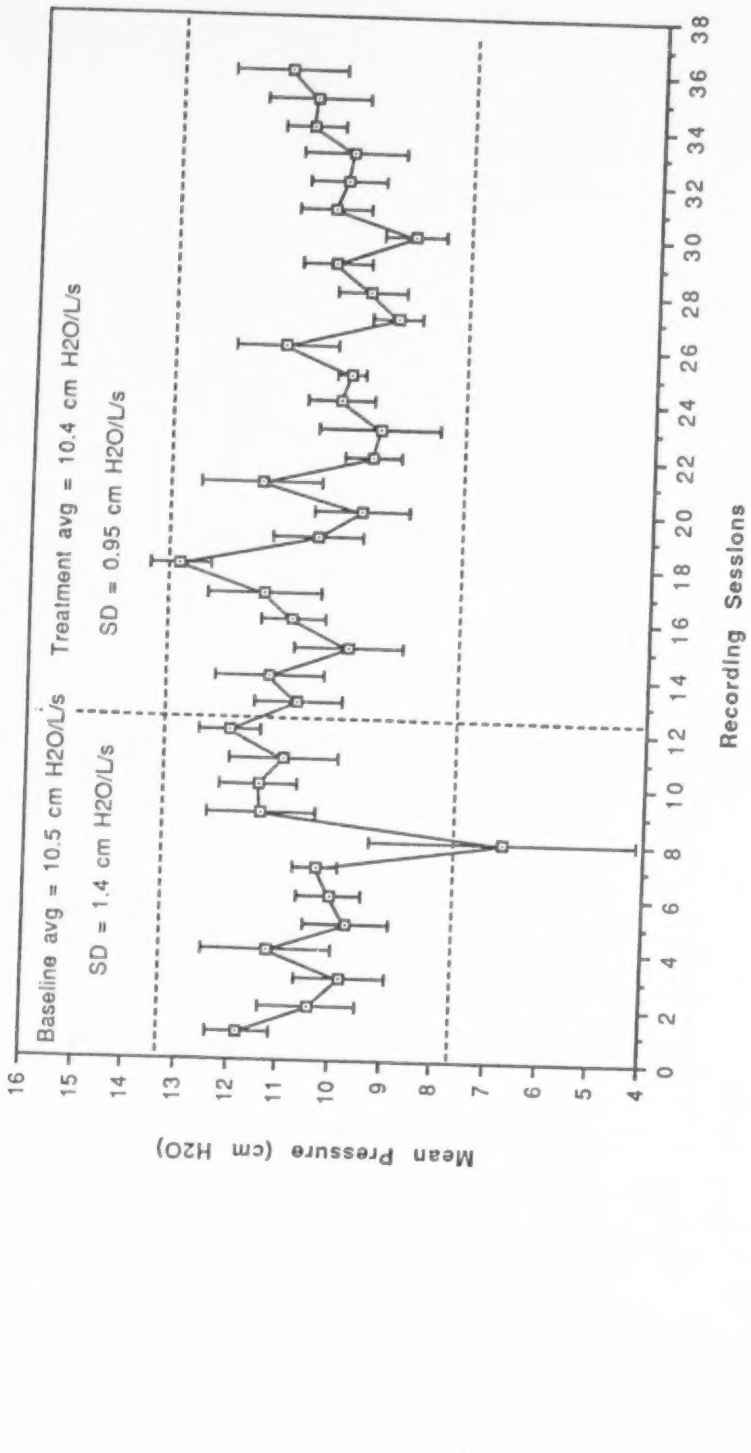


Fig. 4C Subject 2 Flow

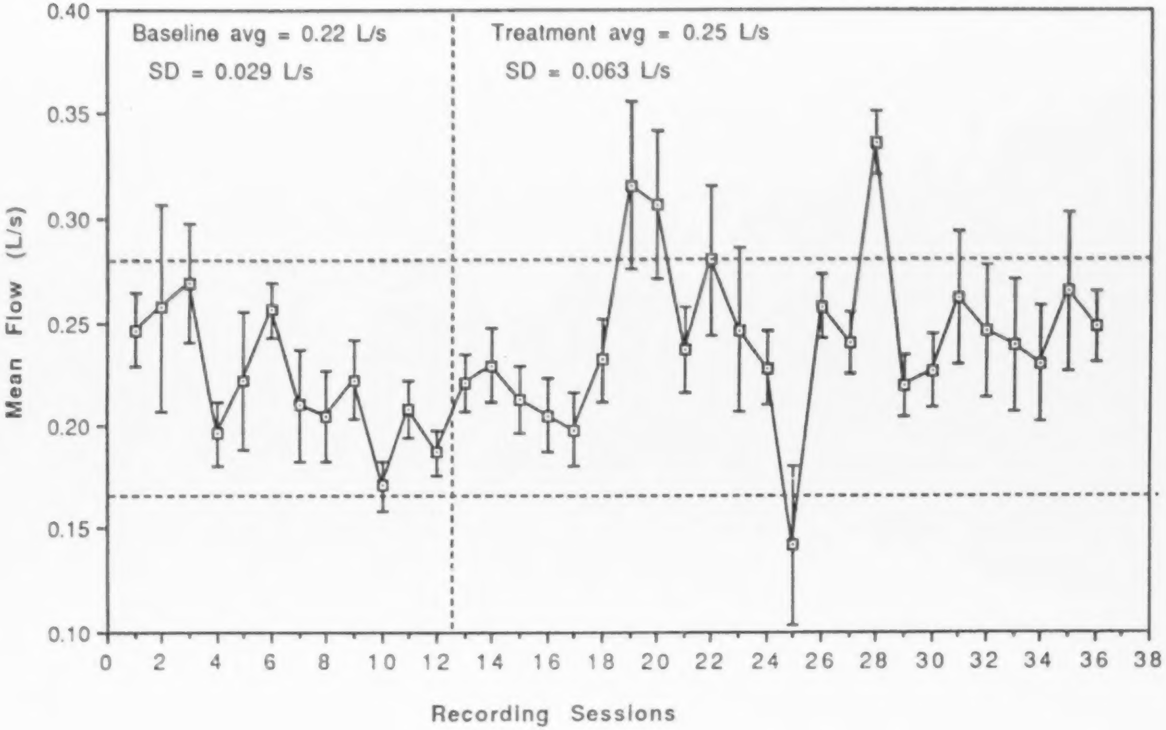


Fig. 5A Subject 3 Resistance

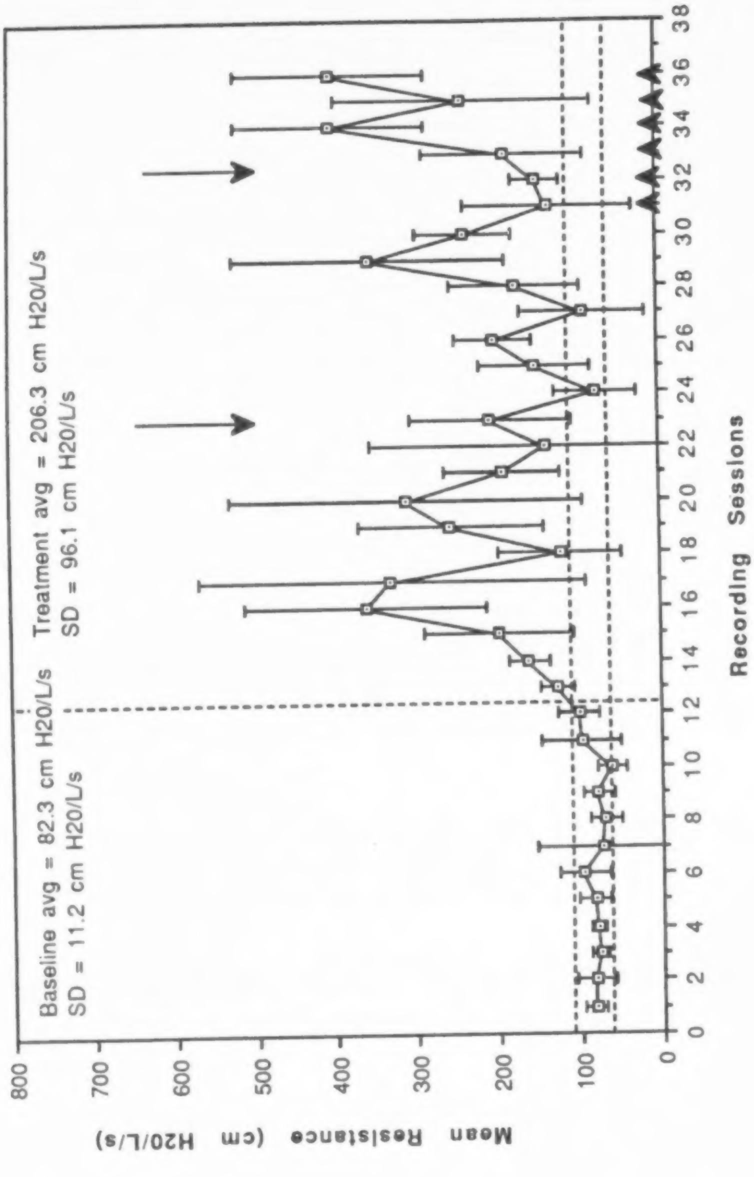


Fig. 5B Subject 3 Pressure

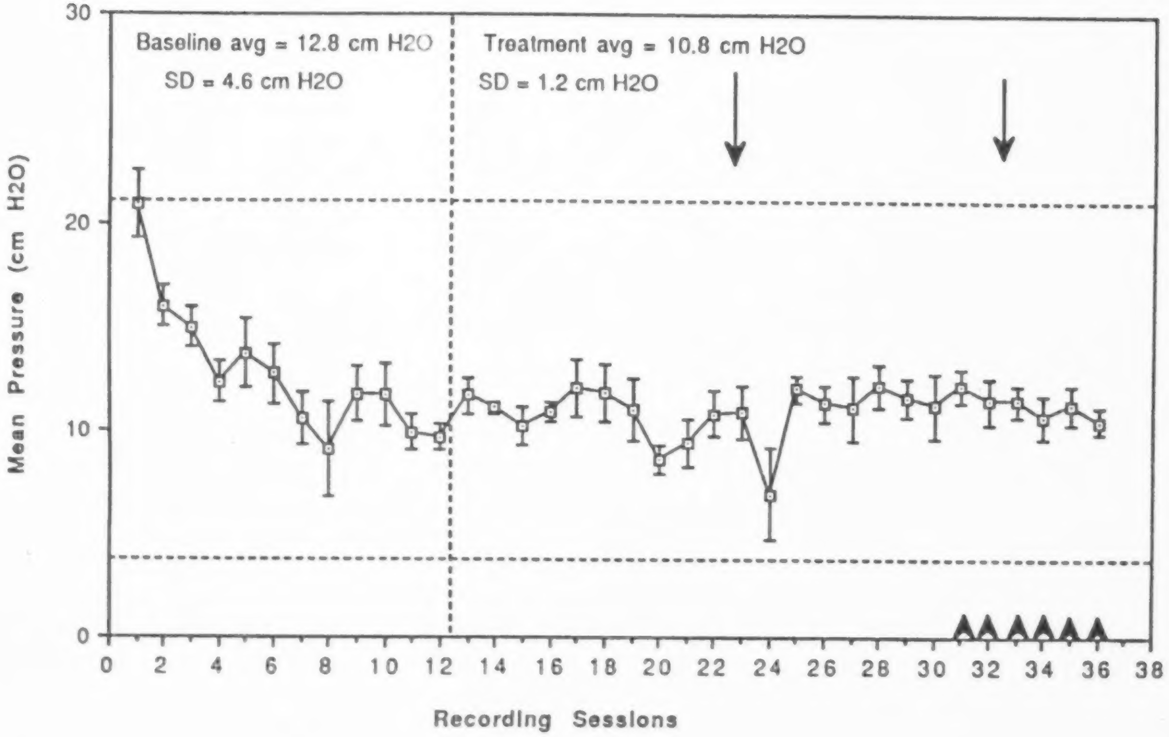


Fig. 5C Subject 3 Flow

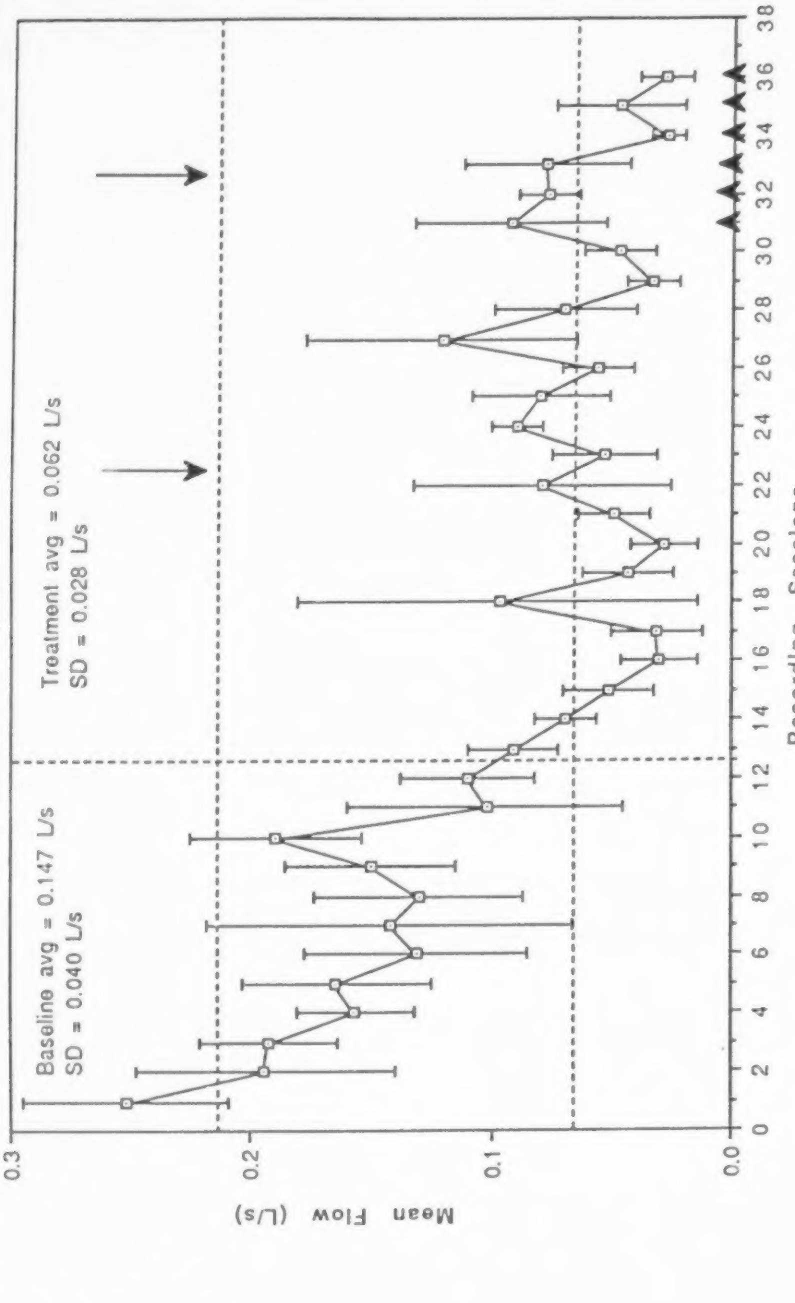


Figure 3A illustrates laryngeal airway resistance data for subject #1 whose mean resistance did not change noticeably across time. Resistance stabilized towards the end of baseline, with the exception of an eccentric data point for the last session (#12). This data point was excluded from calculations of the baseline average printed on the graph. It is related to an extremely low flow value (Fig. 3C) associated with an undetected mask leak. A two-standard-deviation band was constructed with and without the outlying data point for session #12. The wide bandwidth includes the outlier; the narrow bandwidth excludes it. The baseline period average was 34.5 ± 6.9 cm H_2O/L/s, while the treatment period average was 33.0 ± 4.6 cm H_2O/L/s. Within-session variability for resistance was quite small, except for a few sessions at the beginning of the baseline and a few sessions at the beginning of the treatment phase. As shown in Figure 3B, translaryngeal pressure for subject #1 fluctuated within and across her baseline sessions. Variability within and across sessions also was noticeable across the treatment period. The average pressure values for baseline and treatment were virtually the same, however. The baseline average was 11.7 ± 0.92 cm H_2O, while the treatment average was 11.7 ± 1.4 cm H_2O. Three eccentric data points for pressure occurred for sessions #12, #14 and #32. Visual inspection of pressure data points revealed that the low value on session #12 appeared to coincide with a low flow value during the same session. The wide two-standard-deviation bandwidth includes the data point for session #12, while the narrow bandwidth excludes it. An auditory review of the taped session data suggested that the low value on session #14 corresponded with reduced vocal intensity, while the high pressure value on session #32 corresponded with an increase in vocal intensity. As shown in Figure 3C, average translaryngeal flow for this subject's baseline and treatment periods changed minimally from a baseline average of 0.34 ± 0.07 L/s to a treatment average of 0.36 ± 0.06 L/s. Large fluctuations were noticeable in the flow data, however, both within and across sessions in the initial stages of each of the two phases. The outlier for session #12 corresponds to the high resistance value obtained during that session. The wide two-standard-deviation bandwidth includes this outlier, while the narrow bandwidth excludes it.

Figure 4A illustrates laryngeal airway resistance data for subject #2. Her mean resistance did not change sufficiently to exceed the lower limit of the two-standard-deviation band, though there was a trend for resistance to decline across the treatment period. This is reflected in the difference between the resistance baseline average of 48.6 ± 10.3 cm H_2O/L/s and the treatment phase average of 42.0 ± 10.4 cm H_2O/L/, though the stability of this difference is confounded somewhat by large variability both within and between sessions across both phases. An eccentric data point occurred on session #25. The high resistance and the large standard deviation correspond to extremely low flow values that can be attributed to a leak around the face mask. Figure 4B shows that this subject's translaryngeal pressure did not change noticeably across time. The baseline stabilized towards the end of the phase, and the average baseline pressure, 10.5 ± 1.4 cm H_2O, is comparable to the treatment average of 10.4 ± 0.95 cm H_2O. One extremely low data point occurred on the eighth session of baseline and corresponded to a low and extremely variable resistance value for that session. As shown in Figure 4C, translaryngeal flow increased somewhat from a baseline average of 0.22 ± 0.029 L/s to a treatment phase average of 0.25 ± 0.063 L/s. This increase in flow contributed to the decrease in resistance noted for this subject across the experimental period, just as the variability of translaryngeal flow within and between sessions contributed to variance in her resistance data. There were four outliers: one (session #25) occurred below the two-standard-deviation band and corresponded to a high resistance value, and three (sessions #19, 20 and 28) occurred above the the band and corresponded to low resistance values.

Figure 5A illustrates the laryngeal airway resistance data for subject #3 who exhibited a remarkable increase in resistance across treatment, from an average of 82.3 ± 11.2 cm H_2O/L/s in the baseline to an average of 206.3 ± 96.1 cm H_2O/L/s in the treatment phase. The baseline for resistance was very stable. Once treatment began, however, resistance began to increase, and variability both within and between sessions increased remarkably. Twenty-one of the 24 resistance data points are found above the two-standard-deviation band. As shown in Figure 5B, pressure did not change noticeably across treatment once it had stabilized during baseline. At the outset of baseline, pressure values were high. They decreased and stabilized in the latter part of the baseline, which was characterized by an average of 12.8 ± 4.6 cm H_2O, and then remained relatively stable across the treatment phase, averaging 10.8 ± 1.2 cm H_2O. Variability within sessions was small for pressure values. As shown in Figure 5C, flow values for subject #3 decreased significantly from an average of 0.147 ± 0.040 L/s in the baseline phase to an average of 0.062 ± 0.028 L/s in the treatment phase. Flow values exhibited a decreasing trend initially during the baseline. Although they appeared to stabilize during the middle of the baseline phase, they decreased further at the end of baseline and decreased even more during treatment before they appeared to stabilize. Twelve flow data points fell below the two-standard-deviation band, and many of the remaining data points fell along the lower edge of the band. These low flow values correspond to the high resistance values noted earlier. The decrease in flow and the corresponding increase in laryngeal airway resistance are presumed to have been caused by large leaks around the face mask that were undetected by the subject or experimenter.

Menstrual Cycle.

Instances of the reported onset times of the menses are indicated by arrows on the graphs of the aeromechanical data in Figures 3 and 5. Subject #2 was unreliable for these reports, and therefore these data do not appear on her graphs (Figs 4A-C). For subject #3, one onset of menses occurred between the baseline and treatment phases during a week of vacation and is not shown on her graphs (Figs 5A-C). No particular pattern of change was noted in laryngeal airway resistance with the onset of menses for either subject #1 or subject #3, and no reliable menstrual cycle data were available to track for subject #2. In addition, the mid-menstrual cycle periods were examined to note if changes in resistance appeared to be influenced by time of ovulation. Approximate time of ovulation was estimated as the mid-point between two different onsets of menses. No particular pattern of change was noted in laryngeal airway resistance with approximate time of ovulation for either subject #1 or subject #3. No reliable menstrual cycle data were available to track ovulation for subject #2.

Colds.

Subjects reported the onset and duration of colds if any developed throughout the experiment. Two subjects (#1 and #3) developed colds over the course of the experimental period. These are tracked in Figures 3 and 5 illustrating aeromechanical data. Subject #3's cold coincided with the time of menses. Colds did not appear to have a noticeable distinctive effect on either subject's laryngeal airway resistance.

Perceptual Ratings

Order Effect for Paired Sample Presentations.

Before correlations between the dysphonic subjects' laryngeal airway resistance data and listeners' ratings of their voices could be analyzed, the perceptual data sets were examined for evidence of order effects that may have influenced the validity of listeners' judgements in the paired-comparison paradigm. Paired t -Tests (Bruning & Kintz, 1987) to compare the perceptual ratings of each judge between the two orders of presentation (i.e., when a stimulus occupied the first place in the pair versus the second place in the pair) for each subject for both syllable trains and sentences. Results of the t-Tests are presented in Table 2. An order effect was found to be present for two individual raters for only two conditions out of a possible 18. A significant difference between the two presentation orders was found for rater #2 in rating sentences for subject #2, and for rater #3 in rating syllable trains for subject #3. The ratings from these listeners in those two instances were excluded in subsequent data analyses where their inconsistency would confound the results.

Table 2. Intrajudge Order Effects for ratings of voice samples across two orders of paired-comparison presentations.

Subject	Sample	Rater	Order**	Probability
1	Syllables	1	1 vs. 2	0.8749
1	Syllables	2	1 vs. 2	0.5957
1	Syllables	3	1 vs. 2	0.5957
1	Sentences	1	1 vs. 2	0.0858
1	Sentences	2	1 vs. 2	0.8302
1	Sentences	3	1 vs. 2	0.1183
2	Syllables	1	1 vs. 2	0.1728
2	Syllables	2	1 vs. 2	0.2909
2	Syllables	3	1 vs. 2	1.0
2	Sentences	1	1 vs. 2	0.1172
2	Sentences	2	1 vs. 2	0.0145*
2	Sentences	3	1 vs. 2	0.3232
3	Syllables	1	1 vs. 2	0.1067
3	Syllables	2	1 vs. 2	0.4597
3	Syllables	3	1 vs. 2	0.0397*
3	Sentences	1	1 vs. 2	0.1178
3	Sentences	2	1 vs. 2	0.5381
3	Sentences	3	1 vs. 2	0.1829

** Order 1 = The first sample in a paired comparison had been produced earlier in the experimental period than the second. Order 2 = The first sample in a paired comparison had been produced later in the experimental period than the second.

* = significant at 0.05 level. Probability indicated refers to two-tailed probability obtained from paired t-Test.

Trends in Perceptual Data.

Perceptual ratings were tracked by subject and rater over the experimental period in order to determine if listeners perceived voice quality to improve, stay the same or get worse over time . A superficial sample consisting of ratings of 9 sets of paired comparisons for each subject's syllable trains and sentences (one from the fourth week of baseline and one from each treatment week) was used to illustrate trends in the perceptual data. These trends are presented in the form of histograms in Figures 6-8. Figures 6A, 7A, and 8A represent the histograms for the superficial samples of the perceptual trends for the **syllable trains** of subjects #1, 2, and 3, respectively, while Figures 6B, 7B, and 8B represent the histograms for the superficial samples of perceptual trends for the **sentences** produced by the three subjects. The paired sessions that are represented by the ratings are shown on the abscissa of the histograms. In each case the baseline sample (B) was compared with itself and a sample from each week of treatment (T#), based on the listening order in which the baseline recording was presented first and the treatment recording second (a subset of Order #1 in Table 2). A rating of **better** is represented on the ordinate by the numeral **"3"**, a rating of **same** by the numeral **"2",** and a rating of **worse** by the numeral **"1".** Complete illustrations of the perceptual data can be found in Appendix G, figures G-1 to G-6.

For **syllable trains**, raters tended to perceive voice quality at or near the end of treatment to be comparable to or worse than the baseline sample for subject #1 (Figs.6A and G-1), comparable to or worse than baseline for subject #2 (Figs.7A and G-3), and worse than baseline for subject #3 (Figs.8A and G-5). For **sentences**, raters perceived voice quality in treatment to be better than that in baseline for subjects #1 and #2 (Figs. 6B & 7B and G-2 & G-4) and to be worse than that in baseline for subject #3 (Figs. 8B and G-6) .

Figure 6A. Subject 1: Syllable-Trends in Perceptual Ratings.

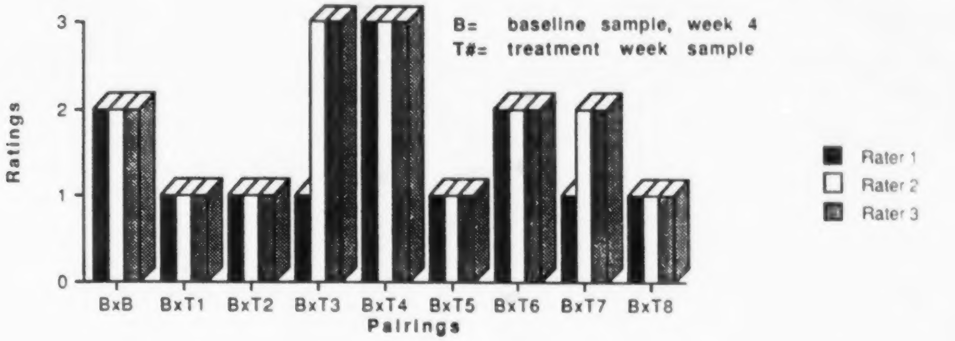


Figure 6B. Subject 1: Sentence-Trends in Perceptual Ratings.

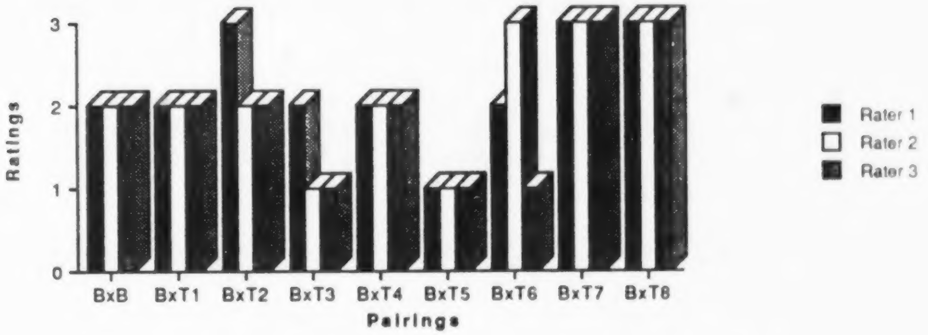


Figure 7A. Subject 2: Syllable-Trends in Perceptual Ratings.

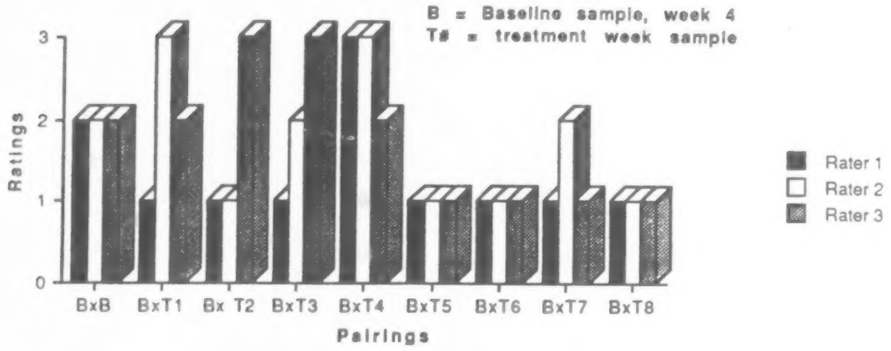


Figure 7B. Subject 2: Sentence-Trends in Perceptual Ratings.

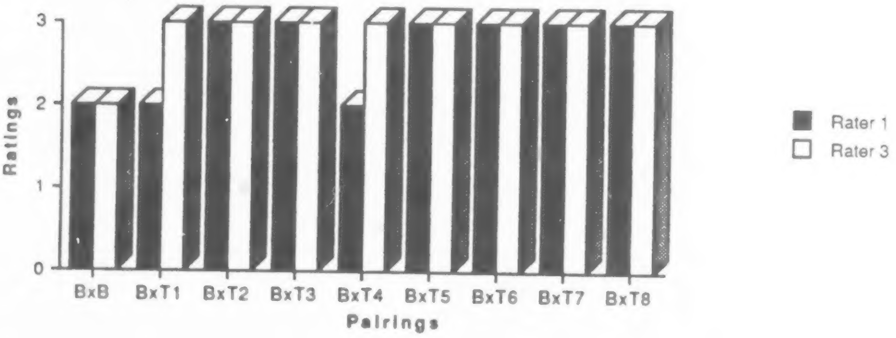


Figure 8A. Subject 3: Syllable-Trends in Perceptual Ratings.

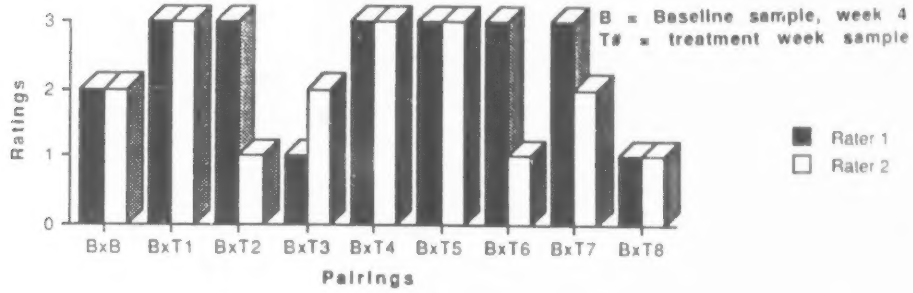
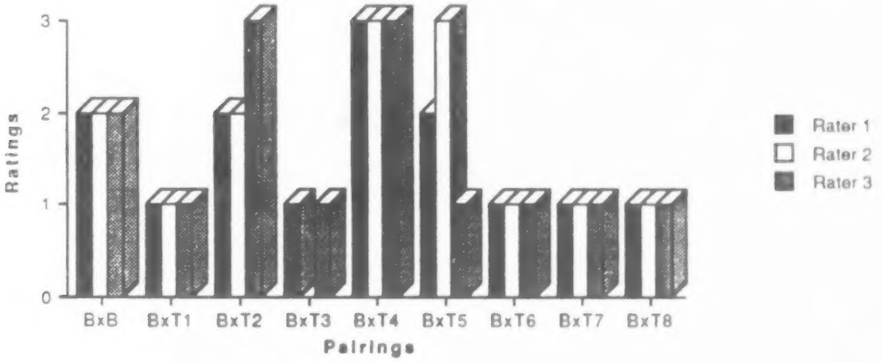

B = Baseline sample, week 4
T# = treatment week sample

Figure 8B. Subject 3: Sentence-Trends in Perceptual Ratings



Specifically, as shown in Figure 6A (see also G-1), for subject #1's **syllables**, raters generally appeared to perceive voice quality to be the same or worse, throughout the experimental period. Some improvement in voice quality was noted in weeks 3 and 4 of the treatment period by most or all of the raters. Beyond this point, however, voice quality for the remainder of the treatment period was primarily judged to be same as or worse than that of the baseline condition. As shown in Figure 6B (see also G-2), for subject #1's **sentences**, voice quality in the treatment samples was generally rated to be the same as or even worse than the baseline sample for treatment weeks 1-6. By treatment weeks 7 and 8, however, all raters generally perceived the voice quality of the treatment samples to be better than that of the baseline recording.

As Figure 7A indicates (see also G-3), for subject #2's **syllable trains**, ratings were quite variable. At the beginning of the treatment period, voice quality was generally rated as same or better than that of baseline by two of the raters. During the last half of the treatment phase, however, voice quality was generally perceived as worse than that of baseline by all raters. As shown in Figure 7B (see also G-4), for subject #2's **sentences**, one rater noted an improvement in voice quality in treatment week 1 in comparison to the baseline recording. By treatment week 2, both raters perceived an improved voice quality that was sustained, with one exception, across treatment. [Ratings for one rater (#2) have been excluded due to an order effect for her results in this condition.]

Figure 8A (and G-5) indicates that for subject #3's **syllable trains**, no consistent rating trends occurred until treatment week 4. During this week and the following one, two raters noted an improved voice quality in relation to the baseline sample. [Data for a third rater (#3) have been excluded due to an order effect for her results in this condition.] After treatment week 5, however, ratings became variable again, and the data displayed for the last session of treatment indicate that voice quality was perceived by both raters to be worse than that of the baseline sample. Figure 8B (see also G-6) indicates that ratings for subject #3's **sentences** also were variable over the course of the treatment period. During the middle weeks of treatment, some raters perceived some improvement in voice quality. During the last weeks of treatment, however, voice samples were rated consistently as worse than the baseline comparator. [Data are missing for Rater 2 in Figures 8B and G-6 for the Baseline vs. Treatment Week 3 comparison as there was no response by this rater for that particular pairing. It is not known how or why such an error occurred in the computer program compiling the response data.]

Intrajudge Agreement.

Table 3 shows the percent exact agreement between each judge's first rating of the voice samples and her second rating of the same samples and order of presentation for each subject. For syllable trains, percentages of exact agreement ranged from 54.3% to 76.3%, while percentages for sentences ranged from 59.5% to 100%. Generally, intrajudge agreement was higher for sentences than for syllable trains, although Rater 1 was more internally consistent when rating syllable trains than sentences for both subject #1 and subject #2, and Rater 3 was more consistent in rating the syllable trains than sentences for subject #2.

Table 3. Intrajudge Percent Exact Agreement for ratings repeated for the same order of paired comparisons.

Rater	Subject Number	Perceptual Sample	% Exact Agreement.
1	1	Syllable Trains	76.3
1	1	Sentences	59.5
1	2	Syllable Trains	71.3
1	2	Sentences	65.4
1	3	Syllable Trains	65
1	3	Sentences	68.8

Overall percentage of exact agreement for Rater 1 = 68%

2	1	Syllable Trains	72.5
2	1	Sentences	74.1
2	2	Syllable Trains	68
2	2	Sentences	75.3
2	3	Syllable Trains	54.3
2	3	Sentences	81.5

Overall percentage exact agreement for Rater 2 = 71%.

3	1	Syllable Trains	67.1
3	1	Sentences	82.3
3	2	Syllable Trains	76
3	2	Sentences	61
3	3	Syllable Trains	63.8
3	3	Sentences	100

Overall percent exact agreement for Rater 3 = 75%.

Interjudge Agreement.

Table 4A shows the percentage of exact agreements among all three judges' ratings for a particular sample and the same order of presentation, which ranged from 34.2% to 67.9.%. Exact agreement percentage scores between pairs of raters (i.e., Rater 1 vs. 2, Rater 1 vs. 3, and Rater 2 vs. 3) are shown on Table 4B. Exact agreements between pairs of raters ranged from 51.3 to 100% for syllable trains, and 47.4% to 69.3% for sentences. Overall chance agreement was calculated to be 33 % for these comparisons between judges.

The Spearman -Rank Order Correlation Coefficient was used to examine the judgements of pairs of raters in order to determine how similar ratings were between them for the same order of presentation. Rho values corrected for ties are shown on Table 4B and range from .031 to 1.0. Twelve of the eighteen correlations were significant at the .05 level and included the following pairs of ratings: Subject #1's syllables and sentences (agreement significant for all pairs of raters), Subject #2's syllables and sentences (agreement significant between Rater 2 and Rater 3 only), subject #3's syllables (agreement significant between Rater 1 and Rater 3 only), and subject #3's sentences (agreement significant for all pairs of raters). Significant correlations exhibited a large range of values but were generally weak to moderate in strength and corresponded to percentages of agreement that were 21.5 to 67 % above the chance level.

Table 4A. Percent Exact Agreement among ratings of all three judges for the same order of paired comparisons.

Subject	**Sample**	**% Exact Agreement.**
1	Syllable Trains	67.9
1	Sentences	44.2
2	Syllable Trains	43.8
2	Sentences	34.2
3	Syllable Trains	40.3
3	Sentences	50.0

Table 4B. Percent Exact Agreement between pairs of raters for the same order of paired comparisons.

Subject	Sample	Pairs of Raters	% Agreement	Rho	Probability
1	Syllables	1 vs. 2	69.1	.455	.0001*
1	Syllables	1 vs. 3	69.1	.455	.0001*
1	Syllables	2 vs. 3	100	1	.0001*
1	Sentences	1 vs. 2	58.4	.268	.0178*
1	Sentences	1 vs. 3	54.5	.251	.0265*
1	Sentences	2 vs. 3	67.5	.531	.0001*
2	Syllables	1 vs. 2	53.8	.155	.1672
2	Syllables	1 vs. 3	51.3	.111	.3227
2	Syllables	2 vs. 3	67.5	.549	.0001*
2	Sentences	1 vs. 2	53.9	.031	.7892
2	Sentences	1 vs. 3	47.4	.041	.7234
2	Sentences	2 vs. 3	61.8	.530	.0001*
3	Syllables	1 vs. 2	52.0	.213	.0595
3	Syllables	1 vs. 3	67.5	.429	.0001*
3	Syllables	2 vs. 3	48.1	.110	.3324
3	Sentences	1 vs. 2	65.3	.344	.0026*
3	Sentences	1 vs. 3	60.0	.303	.0079*
3	Sentences	2 vs. 3	69.3	.447	.0001*

* = significant at 0.05 level. Probability indicated refers to probability values obtained by utilizing a critical-ratio z-test (corrected for ties) to test for significance of rho.

Laryngeal Airway Resistance Correlated with Perceptual Ratings

The primary purpose for obtaining the listeners' perceptual ratings of the recorded voice samples (syllable trains and sentences) was to determine whether there was a relationship between the behavior of subjects' laryngeal airway resistance values across the experimental period and ratings of the subjects' voice quality across that period. The Spearman Rank-Order Correlation Coefficient was used to determine the degree to which differences in laryngeal airway resistance (in cm $H_20/L/s$) correlated with perceived ratings of voice quality. Laryngeal airway resistance difference values were obtained by subtracting the resistance value that corresponded to the second syllable train in a pair from the resistance value that corresponded to the first syllable train in a pair. Samples for comparison came from the fourth week of baseline and each of the eight treatment weeks, and the results of these comparisons are graphed in Figures 9-11. If laryngeal airway resistance values decreased over treatment, relative to baseline, the difference values on the graphs are positive. If laryngeal airway resistance increased with treatment relative to baseline, difference values for paired comparisons are negative.

Table 5 illustrates the correlations between the laryngeal airway resistance difference values and the perceptual rating scores (1=Worse, 2=Same and 3=Better) for the same pairs of syllable trains for all subjects. In addition, perceptual ratings of sentence pairs from the same points in the experimental period were correlated with the laryngeal airway resistance difference scores derived from the syllable trains. The data correlated included comparable samples from the experimental period for all three dysphonic subjects and complete data sets from all three listeners with the exception of the following adjustments neccessitated by experimental error or technological difficulties. First, because the laryngeal airway resistance values for the last session of the baseline period for subject #1 were atypical due to experimental error, the session average resistance value for the second-to-last session of baseline (instead of the last) was utilized in calculating the laryngeal airway resistance difference value for that subject's baseline. In addition, complete sets of 81 pairs of ratings and resistance difference values for each subject's syllable trains and sentences were not always available. A few of the raters' responses were missing in the computerized response logs for some unknown reason; the laryngeal airway resistance difference values that corresponded to these empty cells in the perceptual response sets were eliminated from the pool of values to be correlated. Thus correlations involved 75-80 pairs of resistance difference values and perceptual ratings (instead of 81) from each rater for the syllables and sentences produced by each subject. Furthermore, data were excluded in two instances where an order effect had been identified that threatened the validity of listeners' perceptual ratings (subject #2 sentences by Rater 2, and subject #3 syllables by Rater 3).

Table 5. Correlations between Laryngeal Airway Resistance differences for syllable trains and Perceptual Ratings for the same syllable-train comparisons and for sentences.

Subject	Sample	AVG	Rater 1	Rater 2	Rater 3
1	Syllables	Rho = -.353 p= .0016*	Rho = -.252 p= .0241*	Rho = -.326 p= .0035*	Rho = -.326 p=.0035*
1	Sentences	Rho = .404 p=.0035*	Rho = .123 p=.2854	Rho =.431 p=.0002*	Rho = .313 p=.0063*
2	Syllables	Rho = -.505 p= .0001*	Rho = .039 p=.7256	Rho = -.515 p=.0001*	Rho = -.595 p=.0001*
2	Sentences	Rho = .525 p=.0001*	Rho = .022 p=.8447	**	Rho = .503 p=.0001*
3	Syllables	Rho = .098 p=.3876	Rho = -.032 p=.7796	Rho = .127 p=.2608	**
3	Sentences	Rho = .568 p=.0001*	Rho = .727 p=.0001*	Rho = .322 p=.0047*	Rho = .287 p=.0118*

* = significant at .05. Probability indicated refers to probability values obtained utilizing critical- ratio z-tests (corrected for ties) to test the significance of rho values.

** Rho values and their respective probabilities for these cells have been excluded because order effects threatened the validity of the perceptual judgements of these samples by Raters 2 and 3.

Overall, the correlations shown in Table 5 between perceptual ratings of syllables and corresponding laryngeal resistance difference values and between perceptual ratings of sentences and syllable resistance difference values were weak to moderate in strength. More correlations between sentences and resistance difference values reached significance than did the correlations between syllables and resistance difference values.

For subject #1, there were small but significant **negative** correlations between the perceptual ratings of her **syllable trains** and her corresponding laryngeal airway resistance difference values. The negative correlations are related to a discrepancy between the direction of change in the perceptual ratings and the direction of change in the subject's laryngeal airway resistance across the experimental period. As summarized in Fig 6A, subject #1's voice quality for syllable utterances in treatment generally was perceived to be the same as or worse than her voice quality in baseline with a short-lived exception in treatment weeks 3 and 4. At the same time, as illustrated in Figures 3A and 9, the small decrease in subject #1's laryngeal airway resistance across the treatment period relative to baseline resulted in predominantly positive difference values. Correlations between perceptual ratings of subject #1's **sentences** and her laryngeal resistance values were **positive** and significant for two of the three raters. In this case, Fig 6B shows that raters perceived an improvement in her voice quality on sentence samples by the end of treatment. These trends correlated positively with the predominantly positive differences shown in Fig. 9 that reflect the decrease in the subject's laryngeal airway resistance data with treatment shown in Figure 3A.

For subject #2 there were significant **negative** correlations between perceptual ratings of **syllable trains** by two of the three raters and corresponding resistance difference values. As summarized in Fig. 7A, Raters 2 and 3 perceived the vocal quality of the subject's treatment samples to be the same as or worse than her baseline sample particularly for the last four weeks of therapy. At the same time, Figures 4A and 10 illustrate positive differences that reflect the decrease in subject #2's laryngeal airway resistance data across the treatment period compared to the baseline reference. Correlations between perceptual ratings of subject #2's **sentences** and resistance difference values were **positive** but significant only for Rater 3. [The rho value for Rater 2 was omitted because her ratings for this sample were contaminated by an order effect.] In this case Fig. 7B shows that Rater 3 consistently perceived an improvement in subject #2's voice quality during treatment. These trends correlated positively with the positive differences between baseline and treatment samples shown in Fig. 10 that resulted from a decrease in the subject's laryngeal airway resistance across treatment shown in Fig. 4A.

For subject #3, no significant correlations were found between perceptual ratings of **syllable trains** and corresponding resistance differences. [The rho value for Rater 3 in Table 5 was omitted because her ratings for this sample were contaminated by an order effect.] As summarized in Fig. 8A, Raters 1 and 2 were not consistent in their ratings of subject #3's voice quality during treatment. At the same time, in contrast to the equivocal perceptual trends, Figures 5A and 10 illustrate that negative differences more frequently characterized the increase in the subject's laryngeal airway resistance that occurred during treatment, particularly near the end of therapy. Correlations between laryngeal airway resistance difference values and all three judges' perceptual ratings of voice quality in subject #3's **sentences** were **positive** and significant. In this case, Fig. 8B shows that, with a few exceptions, raters perceived the subject's voice quality in treatment to be the same as or worse than that of her baseline sample. These negative perceptual trends corresponded to the primarily negative difference values shown in Fig.11 between baseline and treatment samples of laryngeal airway resistance as it increased with treatment (Figure 5A).

Figure 9. Subject 1: LAR Difference Values.

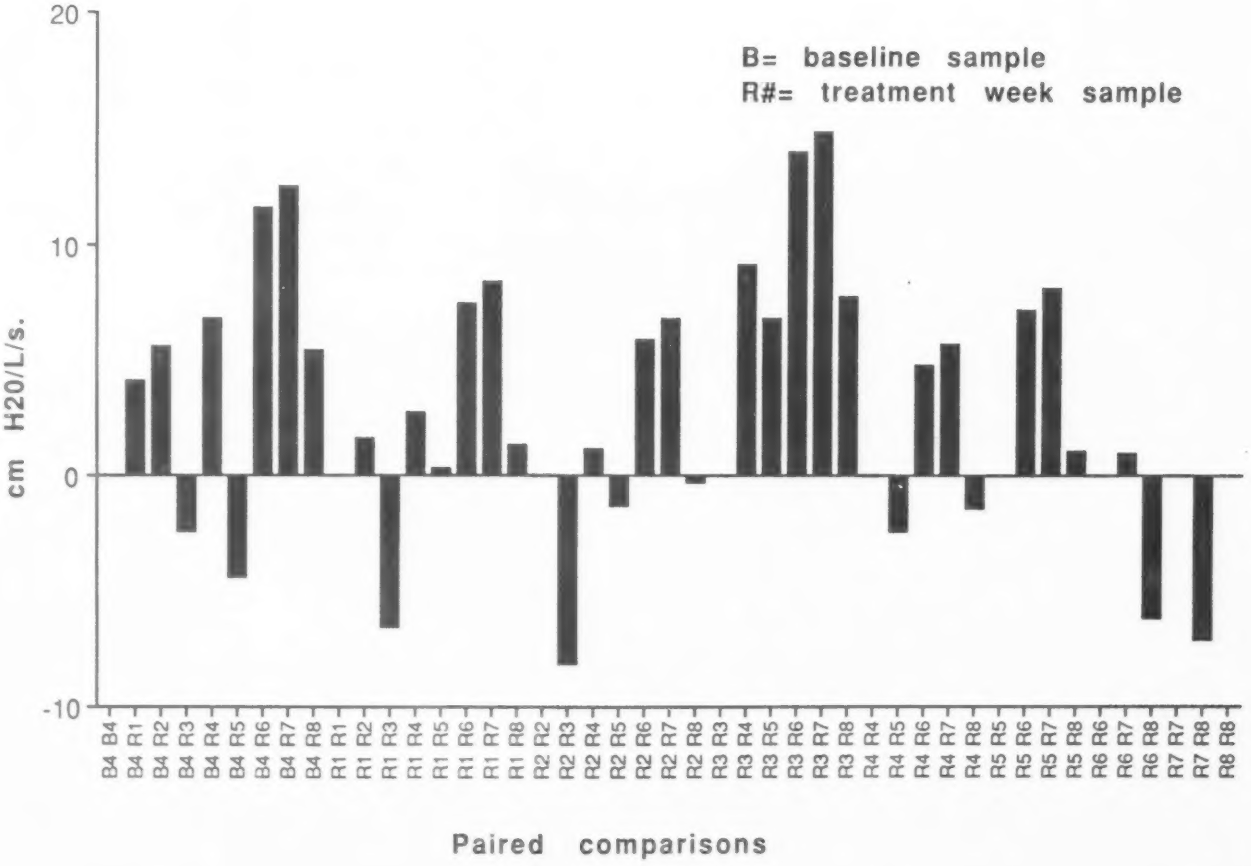


Figure 10. Subject 2: LAR Difference Values.

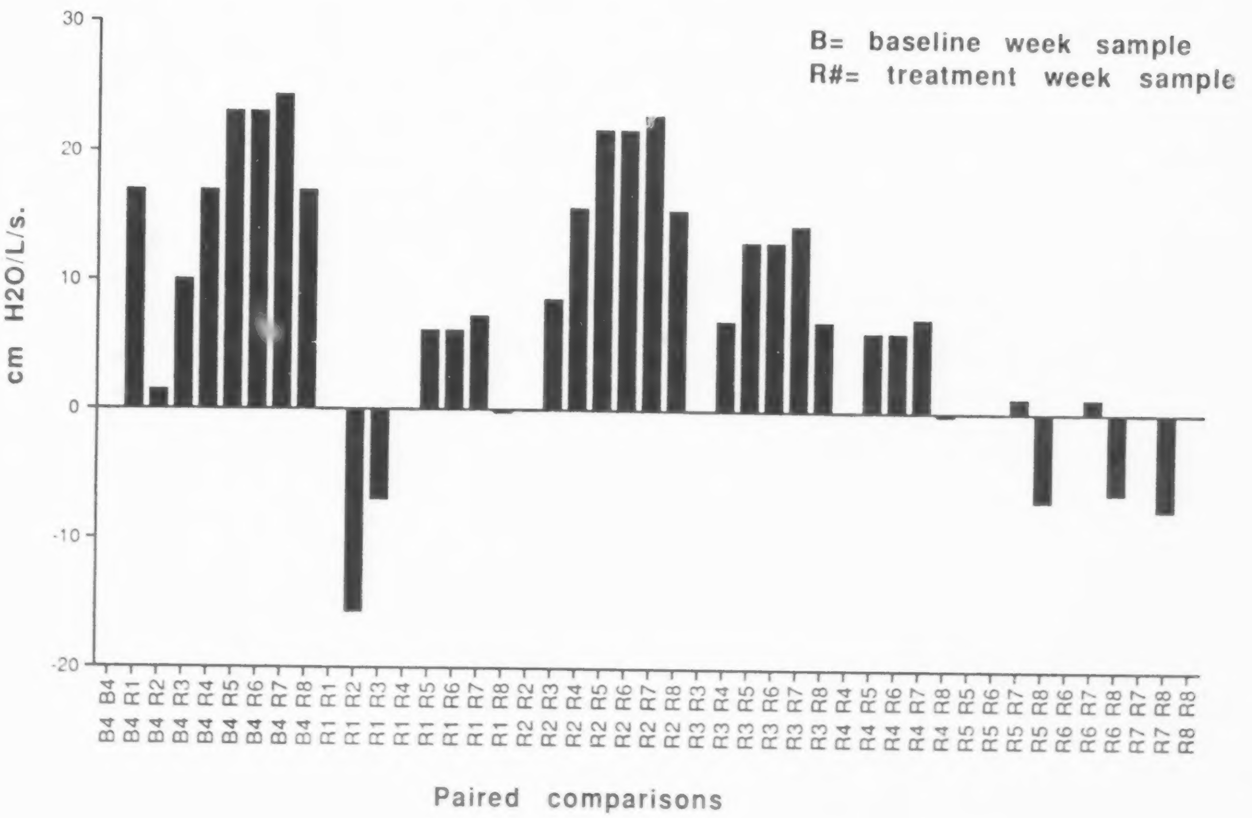
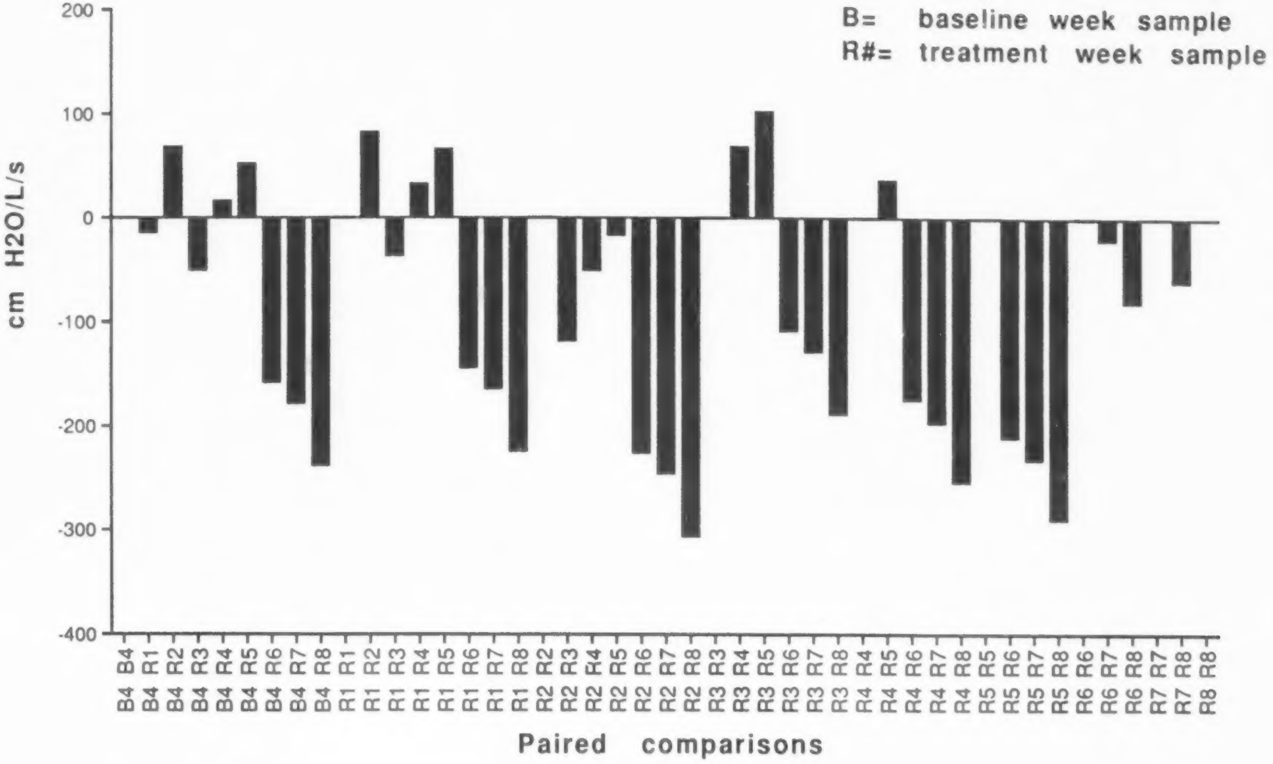


Figure 11. Subject 3: LAR Difference Values.



CHAPTER V. DISCUSSION

The purpose of this research was to observe whether laryngeal airway resistance changed as a result of voice therapy for vocal nodules and whether there was a relationship between aeromechanical and perceptual methods of monitoring changes in the dysphonic voice across a period of treatment. Laryngeal airway resistance measures as well as audio recordings of both syllable repetitions and a reading passage were obtained for three females with vocal nodules across a pretreatment baseline of four weeks and a treatment phase of eight weeks. The syllable trains were the source of the aeromechanical data for calculation of laryngeal airway resistance. The audiorecords of the syllable trains and a sentence from the reading passage from the last week of baseline and each successive week of treatment were used for perceptual judgements. Perceptual ratings of voice quality provided by three clinicians were correlated with changes in laryngeal airway resistance values for each dysphonic subject across the experimental period. A discussion of the findings, clinical implications, and limitations of the present study as well as suggestions for future research will be presented.

Laryngeal Airway Resistance

Research question number one asked whether there was a change in estimates of laryngeal airway resistance in individuals with dysphonia due to vocal nodules across an experimental period of voice therapy. Visual inspection of laryngeal airway resistance data plotted in Figures 3-5 before and throughout voice therapy revealed that this variable remained relatively stable or changed only moderately in two subjects over time and increased remarkably with treatment in a third.

As shown in Figures 3B and 3C, subject #1's pressure and flow values remained fairly stable over the experimental period resulting in a correspondingly stable laryngeal airway resistance (Fig.3A) across time. Several outliers deserve discussion. An extremely low flow value in Figure 3C caused by a mask leak appeared to be responsible for the high resistance value obtained during session #12 (Fig.3A). Another outlier found in session #14 of the pressure data plotted in Figure 3B, contributed to the large variance in laryngeal airway resistance for that session. The low pressure value appears to be explained by reduced vocal intensity heard in the audio record of the same utterance train.

Figures 4B and 4C show that for subject #2, average pressure values remained fairly stable across treatment and average flow values increased slightly. The decrease in the corresponding laryngeal airway resistance value with treatment (Fig.4A) appears to be related to the small increase in the subject's translaryngeal flow during voice production. Two outliers were noted. An outlying low pressure value observed in session #8 (Fig. 4B) appears to be responsible for a correspondingly low and variable resistance value on that day. A low flow value caused by a mask leak (Fig. 4C) is responsible for the high resistance value and its large variance in session #25 (Fig. 4A).

Figures 5B and 5C show the pressure and flow values for subject #3. For this subject, pressure values remained fairly stable across the experimental period (Fig.5B) while flow values decreased remarkably with treatment (Fig.5C). As a result of the decrease in flow values, an extraordinary increase in corresponding laryngeal airway resistance values (Fig.5A) occurred with treatment. These low flow and high resistance values are unrealistic even for an extremely dysphonic patient and are likely the artifact of an experimental error. Apparently airflow leakage around the facial mask occurred on data collection throughout the experimental period for this subject but was not detected by the experimenter.

Thus, based upon visual inspection of laryngeal airway resistance data illustrated in Figures 3-5, it appears that laryngeal airway resistance did not change in a reliable manner with treatment: for subject #1, resistance values remained stable over the experimental period, for subject #2 resistance decreased slightly with treatment, and for subject #3, resistance increased dramatically.

It is informative to compare the laryngeal airway resistance results of the subjects in the present study with the data reported in earlier research. Several researchers have provided reference data for laryngeal resistance in young women. For example Leeper and Graves (1984) reported laryngeal airway resistances for 15 normal young adult females between the ages of 23- 32 years under controlled-intensity experimental conditions. The overall group mean for laryngeal airway resistance was 38.3 cm H_2O/L/s with a standard deviation of 9.25 cm H_2O/L/s. Hoit and Hixon (1992) reported a mean laryngeal resistance value of 54.88 cm H_2O/L/s with a standard deviation of 15.04 cm H_20/L/s. in an uncontrolled-intensity condition for female adults aged 25-35 years.

Because the present study was modelled after that of Leeper & Graves, it will be discussed in reference to their results. Subject #1's average resistance for baseline (34.5 ± 6.9 cm H_2O/L/s) and for treatment (33.0 ± 4.6 cm H_2O/L/s) fell within one standard deviation of Leeper & Graves' group mean of 38.3 (± 9.25) cm H_2O/L/s. Subject #2 may not be comparable to Leeper and Graves' subjects because she was younger than their group by at least 8 years. Her average baseline resistance (48.6 ± 10.3 cm H_2O/L/s) was slightly beyond one standard deviation above the Leeper and Graves' group mean and remained so at the close of treatment (45.4 ± 13.6 cm H_2O/L/s). For subject #3, resistance values were considerably higher during baseline than what would be expected for adult females (82.3 ± 11.2 cm H_2O/L/s) and became extremely and unnaturally higher during treatment (206.3 ± 96.1 cm H_2O/L/s). As noted earlier, these extremely high resistance values appear to have been caused by a leak around the face mask. This artifact precludes legitimate comparison of subject #3's data with those of Leeper and Graves' subjects.

The behavior of laryngeal airway resistance with treatment was also examined with respect to the structural changes documented in the subjects' vocal nodules. No reliable relationship was found between these factors, however. Subject #1's laryngeal airway resistance remained stable across the experimental period, while her vocal nodules disappeared with treatment, though her vocal folds appeared somewhat edematous at the close of treatment. For this subject, laryngeal airway resistance did not appear to reflect the change that occurred in her nodules. Subject #2's laryngeal airway resistance decreased slightly across the experimental phase and her vocal nodules also disappeared by the close of treatment. For this subject, a slight decrease in laryngeal airway resistance appeared to parallel the structural changes that occurred in the vocal nodules. Subject #3's laryngeal airway resistance increased noticeably from baseline to the end of treatment, while her vocal nodules reduced in size from approximately 2-3 mm at baseline to less than 1mm at the close of treatment. For this subject, an extraordinary increase in laryngeal airway resistance values paralleled the change observed in the status of her vocal nodules.

The absence of an apparent relationship between changes in laryngeal airway resistance and all three subjects' physical responses to voice therapy (i.e. reduction in or elimination of their vocal nodules) may be related to several threats to the internal validity of this study.

Instrumental and Procedural Errors.

Several experimental errors in the present study constitute threats to the internal validity of the laryngeal airway resistance data as indicators of a response to voice therapy. One of the most influential procedural errors was the undetected presence of flow leaks around the edges of the mask where it rested against the subject's face. These flow leaks contributed to low flow values and consequently high resistance values and large standard deviation for both parameters. Such flow leaks occurred on a few occasions for subjects #1 and #2 and occurred throughout the treatment period for subject #3. The fact that a mask leak was undetected for such a substantial portion of her treatment period, essentially invalidates subject #3's laryngeal airway resistance data for the experimental question of interest.

Faulty operation of the sound-level meter constitutes a second instrumental problem that reduced the reliability of intensity control such that high pressure outliers and large variability in pressure values were recorded on some occasions in all subjects. After the end of the treatment period for the last subject it was noted that the batteries in the sound-level meter were low. It is not known when this problem began, but the low voltage may have affected the sound-level meter's performance, thereby allowing inconsistent intensity levels from session to session.

A third procedural oversight that affected the reliability of the data obtained for laryngeal airway resistance is the fact that subjects (especially subject #1) often produced the /pi/ syllable fewer or more than seven times on one breath group during the laryngeal airway resistance task. This inconsistency meant that subjects were not always working in the same lung volume range across sessions which could have influenced laryngeal airway resistance and its variability. In addition when subjects produced fewer than five syllables per breath group, data analysis was limited to fewer useable samples.

Vocal Intensity.

Variance in the vocal intensity utilized by subjects in the laryngeal airway resistance task was a potential threat to the internal validity of the data obtained as indicators of change related to voice treatment. Leeper and Graves (1984) had found laryngeal airway resistance to be more stable and standard deviation figures to be reduced substantially when intensity was controlled at a level of approximately 75 dB. This study had been modelled after theirs to include intensity controls in subjects' utterances during aeromechanical data collection. The subjects in this study were required to maintain their intensity between 64 and 70 dB while recording the syllable trains used for calculation of laryngeal airway resistance. This range of intensity was chosen on the basis of a pilot study that identified comfortable performance levels for loudness during such data recording. In contrast to Leeper and Graves' study, in which the intensity control was thought to have a stabilizing effect on the variance in the aeromechanical data, the present study's data exhibited relatively large standard deviations for laryngeal airway resistance for two of the three subjects. It is likely that battery problems with the sound-level meter contributed to some of the variance observed. This in turn, limits the extent to which changes in laryngeal airway resistance can be interpreted as related to subjects' responses to voice therapy.

Baseline Stability.

The present study employed an A-B single-subject design to provide an opportunity for laryngeal airway resistance measures to stabilize in the dysphonic subjects before the onset of voice therapy. Under ideal experimental conditions a stable baseline facilitates the evaluation and interpretation of changes that occur during the subsequent treatment period (McReynolds & Kearns, 1983). Subject #1's resistance values in baseline stabilized towards the end of the pre-treatment period with the exception of one outlier which ocurred on the last session of this phase. Subject #3's resistance values remained fairly stable across the baseline phase. On the otherhand, the baseline established for subject #2's resistance was variable. Variability in flow values, which is likely attributable to mask leaks, appears to account for the lack of stability in the baseline of resistance for this subject. Baseline instability in laryngeal airway resistance diminishes the confidence with which the change in resistance exhibited by subject #2 may be interpreted as a response to treatment and widens the range that the treatment data must exceed in order to represent a significant change.

Menstrual Cycle.

The experimenter attempted to track the menstrual cycle in the subjects of this study in order to determine if hormonal influences were a possible threat to the internal validity of the laryngeal airway resistance data. Endocrine changes during menstruation may cause vocal folds to become edematous as a result of water retention (Greene & Mathieson, 1991). As a result, voices may be dysphonic in the premenstrual and menstrual periods (Flach, Schwickardi & Simon, 1969). Onset of menses during the experimental period was tracked and plotted on Figures 3 and 5 for subject #1 and subject #3. Subject #2's reports of onset of menses were unreliable and thus could not plotted. Visual inspection of Figures 3 and 5 reveals no obvious fluctuations in laryngeal airway resistance that appeared to coincide with onset of menses.

Higgins and Saxman (1989) found increased variation in jitter at the time of ovulation and speculated that when ovarian hormones fluctuate, the vibratory stability of the vocal folds is affected. These researchers suggested that the influence of these hormones may be strong enough to influence the voice of a non-singer during speaking. The possible influence of ovulation on laryngeal airway resistance for subjects #1 and #3 was inferred from observations of their mid-cycle resistance data. Visual inspection of those data in Figures 3 and 5 also revealed no noticeable temporary change in resistance corresponding to the estimated time of ovulation.

Colds.

The effect of colds on laryngeal airway resistance was also a potential threat to the internal validity of laryngeal airway resistance data as an indicator of vocal improvement. During a severe cold accompanied by laryngitis, a patient's voice may be lower in pitch and hoarse due to severe swelling and redness of the membrane that covers the vocal folds (Boone & McFarlane, 1988). The experimenter attempted to monitor this potential threat by tracking subjects' cold symptoms across the experimental period. Both subject #1 and subject #3 developed colds over the course of their participation in this study, and the onset and duration of symptoms are documented in Figures 3 and 5. Visual inspection of the data collected at these points in the context of neighbouring data, suggests that colds did not appear to have a noticeable effect on either subject's laryngeal airway resistance.

Summary of correspondence between LAR and treatment outcome.

The data obtained for laryngeal airway resistance in this experiment did not change in the same direction or to the same degree with treatment across the three subjects, nor did the unique changes that occured appear to be sensitive to changes in the status of subjects' vocal nodules with treatment. Although the nodules disappeared entirely in two subjects and resolved noticeably in the third, corresponding laryngeal airway data remained essentially unchanged, decreased moderately or increased dramatically with treatment. These results cannot be used as evidence that laryngeal airway resistance may not change reliably with voice therapy, however, because the validity of the resistance data was compromised by significant instrumental and procedural errors, and possibly by baseline instability in some or all of the subjects.

Correlations between Aeromechanical and Perceptual Measures

Research question two explored whether there was a relationship between estimates of laryngeal airway resistance and perceptual judgements of voice samples obtained from the three subjects with vocal fold nodules across an experimental period of voice therapy. Aerodynamic and perceptual measures representing data from the last week of the baseline phase through the last (eighth) week of the treatment phase served as the basis for this comparison. The results of these correlations were variable in strength and sign, and even those that reached significance revealed no reliable relationship between the behavior of subjects' laryngeal resistance across the experimental period and listeners' perceptions of the subjects' voices across the same period.

The reader will recall that subtraction of LAR values obtained during treatment from a representative baseline value yielded single difference values whose sign and magnitude reflected the behavior of LAR across the experimental period. These difference values for all possible paired comparisons in a single presentation order are plotted in Figures 9, 10 and 11. If LAR had decreased with treatment (which happened slightly for Subject 1 and moderately for Subject 2), difference values exhibited a positive sign. If LAR had increased with treatment (which occurred dramatically for Subject 3) the difference values exhibited a negative sign. These difference values for resistance were correlated with the perceptual judgements of three clinicians for the same voice sample comparisons from which the LAR difference data were derived. Clinicians had been asked to rate the second syllable-train sample in each pair as the same as ("2"), worse than ("1") or better than ("3") the first sample. In addition, the listeners rated comparisons of the sentences produced by the subjects for the same sessions across the experimental period, though the sentences were not related physiologically to the LAR data. Listeners' ratings for all possible paired comparisons in a single presentation order are plotted in Figures G1 - G6.

Table 5 illustrates the correlations obtained between the laryngeal airway resistance difference values and the perceptual ratings of the syllable trains and sentences for all subjects. Significant negative correlations characterized the relationship between laryngeal resistance and perceptual ratings assigned to **syllable-sample** comparisons for subjects #1 and #2. Perceptual ratings of the syllable data were "the same as" or "poorer than" baseline with treatment (i.e., ratings of "2" or "1" predominated). At the same time, laryngeal airway resistance in these two subjects decreased with treatment, more or less, resulting in difference values that were positive in sign. Lower ratings correlated with LAR differences that were positive in sign may have contributed to the opportunity for correlations to be negative, an artifact of the nature in which the laryngeal difference values were derived. A clinical explanation also may be offered to interpret these significant negative correlations. Physiologically, a reduction in laryngeal airway resistance is related to an increase in translaryngeal flow during voicing, which may be perceived as breathiness. "Breathiness" was among the adjectives used to describe a dysphonic quality for the listeners who rated the voices. Thus, if decreased LAR was associated with an increased in perceptible breathiness in these two subjects, it is not surprising that listeners rated their voices as poorer with treatment, even though a decrease in LAR with treatment for nodules may also be accompanied by a decrease in laryngeal tension and other hyperfunctional behaviors associated with the develoment of nodules.

The correlations between the perceptual ratings assigned to **sentence-sample** comparisons and LAR that reached significance were all positive. Two raters perceived the voice of subject #1 in the sentences to improve with treatment; higher ratings correlated positively with LAR difference values that were positive in sign. One rater perceived the voice of subject #2 in sentences to improve with treatment. Once again, higher ratings correlated positively with this subject's positive LAR difference values that reflected the moderate decrease in her resistance with treatment. For subject #3 correlations also were positive for sentences and LAR data, though her data behaved in a manner opposite to that for subjects #1 and #2. Her sentence ratings got poorer with treatment while her LAR increased dramatically, a change that was reflected in negative difference values. In all these cases for the sentence material, the interpretation of the signs of the significant correlations is shrouded in even more ambiguity than the interpretation of the correlations between ratings of syllable trains and LAR. Certainly the signs of the correlations may be an artifact of the method by which the laryngeal difference values were obtained. Listeners also may have been influenced by more than just breathiness in the sentence material, rendering a physiological interpretation of the voice quality ratings moot. And finally, it must be remembered that the basis of a significant correlation is purely speculative in these cases, because the sentence data are not related physiologically to the LAR data.

The correlations between aerodynamic and perceptual measures for the sentence material in this research may be compared cautiously to two other studies that attempted to study the relationship between laryngeal valving during voice production and listeners' perceptions of the quality of the vocal output. Harris (1971) and Moran and Gilbert (1984) collected data about translaryngeal airflow during the recording of connected speech samples that were then rated for quality. They reported Spearman Rank Order Correlation coefficients of 0.56 (Harris, 1971) and 0.44-0.72 (Moran & Gilbert) between their flow data and perceptual ratings. The range of these correlations is comparable to those achieved in the present study, but their apparent similarity must be interpreted with caution for two reasons: (1) The aeromechanical and perceptual variables correlated in the previous studies were derived from the *same* speech sample, whereas the present study compared LAR data derived from one sample (syllable repetitions) to perceptual data for another (sentence utterance); and (2) the previous studies correlated only average flow values with perceptual ratings for the speech samples, whereas the present study correlated LAR (estimated from flow *and* pressure) with perceptual rating data.

The mixed and ambiguous results of the correlations between laryngeal airway resistance and perception in the present study likely reflect two major sources of experimental error. Threats to the validity of the laryngeal airway resistance data already have been discussed. The validity of the perceptual data used in the comparisons also may have been threatened by procedural and psychophysical phenomena, and these will be considered in the next section.

Procedural threats to the validity of the perceptual data.

Procedural problems that may have influenced the audio data collection include artifacts or errors in the voice-sample recording process, some of which have been mentioned already in the discussion of data acquisition for laryngeal resistance. Artifacts produced in the recording process of syllable trains may have affected raters' perceptual judgements. In two instances (Subject #2, Treatment Week 5 and Treatment Week 3) where quality of the recording was compromised considerably, the distorted syllable trains were replaced by other syllable samples produced during the same treatment session. A number of signal artifacts --for example, distortions in which plosives were voiced (/b/ instead of /p/) and prosody was altered -- occurred during the recording for a small number of the syllable trains of subject #3 which could not be replaced or repaired. Although the listeners were reminded to rate voice quality only on the basis of diplophonia and breathiness, they reported that these signal distortions made ratings more difficult.

In addition to sound artifacts or distortions that may have been included on the voice recordings, it is acknowledged that the recording process itself influenced the quality of the signals rated by the listeners. Because the syllable-train samples served as a source for aeromechanical as well as acoustical data recording, the apparatus associated with the aeromechanical data collection may have influenced subjects' production of the syllables as well as the spectrum of the signal that was recorded on tape. This influence included primarily the effects of the full-face mask and the airflow collection system (mask plus pneumotachometer). The mask damped the intensity of the voice signal as it was produced by the subjects, and the flow measurement system increased the resistance subjects had to overcome during the production of voice. Both of these influences might have encouraged subjects to use more effort during voice production, to make the voice signal more audible to themselves and to compensate for the additional downstream resistance of the aeromechanical instruments. Furthermore, because the recording microphone was sampling outside of the mask and pneumotachometer, the spectrum of the vocal signal was no doubt filtered by the flow-collection assembly. While all these influences are acknowledged as inevitable in instrumental arrays that sample aeromechanical and acoustical data simultaneously, their effects on the subjects' vocal output should have been constant across the experimental period thus limiting them as uncontrolled threats to the validity of the data recorded. The one procedural exception that threatened this experimental control was inconsistency in loudness across the recordings. Although the experimenter had intended to control intensity across the experimental period, faulty operation of the sound level meter reduced the reliability of intensity control such that high pressure outliers and large variability in pressure values were recorded on some occasions in all subjects. Raters were asked to make their perceptual judgements of voice quality based on breathiness or diplophonia. Even so, changes in the intensity of the voice samples may have influenced the perceptual gestalt of vocal quality and thereby have biased listeners' judgements, although the raters did not report this to be a problem for them.

Psychophysical threats to the validity of the data used in the correlations.

Psychophysical phenomena related to the information contained in the signals rated by the listeners and the nature of the rating process itself also may have influenced the perceptual results and their correlations with laryngeal resistance.

The nature of the perceptual sample. The limited sample size of the syllable-train data may have influenced listeners' ability to rate them reliably in the case of subject #2 for whom only three syllables/train were available for rating because of errors in data collection. A more serious limitation of the syllable trains for perceptual ratings for all subjects is related to the information about voice quality contained in such a sample. The repetition of /pi/ on one expiratory breath as performed by the speakers in this study suffices for laryngeal airway resistance calculations but provides a listener with an extremely limited sample for voice quality judgements, namely several opportunities to perceive voice onset and then a short term of sustained phonation followed by voice offset. This is hardly comparable to the sample of sustained voice and connected speech usually required in clinical practice to rate voice quality with respect to breathiness, diplophonia and a number of related characteristics. While standard clinical practice in voice therapy might include a short recording of a vowel prolongation produced on one expiration, the primary measure of voice quality comparison includes a connected speech sample which usually consists of a one-paragraph recording of a standard reading passage (Prater & Swift, 1984; Greene & Mathieson, 1991; Colton & Casper, 1990) Thus, while it was important to sample listeners' judgements of the same vocal signals that served as the source of laryngeal resistance data, it is acknowledged that the aeromechanical sample and the acoustical sample of the same utterance were not equal in their ability to characterize the vocal features of interest in each parameter. It is for this reason that short samples of connected speech in the form of a sentence also were included as part of the rating process. Although the sentence data could not be analyzed for information about laryngeal airway resistance, they were useful in providing some insight into the reliability of listeners' ratings for the syllables versus the sentences and ultimately, perhaps, the validity of correlations between the perceptual data and laryngeal resistance. As discussed below along with other issues related to listener reliability, the fact that intrajudge agreement tended to be higher for sentence samples than for syllable trains (Table 3) may be indirect evidence that the syllable train data were inadequate for reliable perceptual assessment. This observation and all of the perceptual results must be considered in light of literature that suggests listeners are not necessarily reliable in rating voice quality, even under the best of listening conditions and accumulated experience.

Listener reliability and the nature of the listening task. Intrajudge reliability and interjudge reliability were variable and rarely high. The overall chance agreement had been calculated to be 33 %, and thus intrajudge reliability for syllable trains (Table 3), which ranged from 54-76 %, was not very high. Intrajudge reliability for sentences (also Table 3) exhibited a large range (59-100 %) that suggested listeners were more decisive.

Intrajudge reliability across all subjects (Table 3) was highly similar (68%, 71%, and 75% for Raters 1, 2, and 3 respectively), and intrajudge reliability within each subject was also very similar (72 % for subjects #1 and #3 and 70 % for subject #2). Subject #2's intrajudge reliability was the lowest. Raters commented that her voice samples were the most difficult to rate. Subject #2's nodules were not well defined, and it was the investigator's impression that her voice quality was more mildly dysphonic in relation to the other two subjects. The fact that this subject's dysphonia was not very severe to begin with, may have made any change in quality more difficult to detect.

Percent exact agreement among and between raters (Tables 4A and 4B) was also quite variable. When compared to the chance level (33%), interjudge exact agreement percentages between pairs of raters (Table 4B) was also not very high (54.5 to 100 %). The Spearman-Rho Correlation Coefficient was used to correlate ratings between pairs of raters (Table 4B) and yielded weak to moderate correlations most of which nevertheless reached significance.

The fact that both intra-and interjudge agreement data were variable and only moderately reliable renders questionable the validity of the perceptual data used in the correlations analyses. However, when intra- and interjudge agreement data are examined in the context of other studies of this nature, the data are consistent with recent reports that have examined the complex phenomenon of voice quality rating (Kreiman et al, 1993).

Previous studies of voice quality judgements (Deal & Emmanuel, 1978; Netsell, Lotz & Shaughnessy, 1984; Yumoto, Sasaki & Okamura, 1984; Bassich & Ludlow, 1986; and those of Kreiman et al., 1990a, 1990b, 1992, 1993) reported listener reliability results fairly similar to those in the present study. The most recent study by Kreiman et al. (1993) involved 20 clinicians who rated roughness in voice quality and obtained an overall intrajudge percent exact agreement of 47.5 % with a range of 20-63%. Kreiman et al.

(1993) reported a mean exact agreement among raters of 33.7%, with a range of 6.7%-56.7%.

Kreiman et al. (1993) reviewed the literature of past perceptual studies using various protocols and reported that intra- and interjudge reliability did not vary consistently with level of task-specific training or the actual rating task used. Thus Kreiman and colleagues argued that even "highly experienced listeners frequently disagree completely about what they hear" (p.33). They suggested that perfect agreement among raters is an impossibility even in theory. Nevertheless, these researchers offered a solution to the perceptual problem. They suggested that "variability in voice quality ratings might be reduced by replacing listeners' idiosyncratic, unstable, internal standards with fixed external standards or "reference voices" for different vocal qualities (p.33). It was this suggestion that influenced the use of paired comparisons as a method of perceptual judgement in the present study, in hopes that reliability would be improved when two perceptual samples were presented one after another. In this format when a rater makes a perceptual judgment about one voice sample she can compare it to a relevant external reference point (the other perceptual sample) rather than against an internal standard that may be different from the standards of the other judges. However, the results of this study showed that raters require more than just paired comparisons to make reliable perceptual judgements. Future studies of voice perception must consider not only the format in which voice samples are presented for evaluation but also other factors that may influence the perceptual task: namely, (1) procedural factors related to the recording of the samples, (2) the nature and duration of the sample chosen for evaluation, (3) the variability of the perceptual dimensions of interest across samples, and (4) the ability of listeners to identify and agree upon those perceptual dimensions.

Summary of correlations between LAR and perceptual data.

Although significant correlations occurred between data representing changes in subjects' laryngeal resistance with voice treatment and listeners' perceptions of those subjects' voices across the treatment period, the validity of the correlations is rendered ambiguous by a number of procedural, sampling and psychophysical factors. Artifacts and errors in the voice recordings, particularly with respect to loudness, may have affected the consistency of the recordings and influenced listeners' perceptions of them. The syllable samples obtained for aeromechanical and perceptual analyses were adequate for the former but probably not for the latter, while the sentence samples appear to have been more informative for perceptual analysis but were inadmissable for aeromechanical analysis. Finally, intra- and interjudge agreement data were variable and only moderately reliable, which is consistent with reports in the literature for similar rating tasks and renders questionable the validity of the perceptual data used in the correlational analyses.

External Validity Issues

External validity issues also limit the extent to which the results of this study can be generalized to clinical evaluations and treatment plans for other subjects with vocal nodules. Sample size was small (n=3), and the behaviors of the dependent variables across subjects were inconsistent. Thus, the use of only three subjects in an A-B single-subject-design with no within-subject replication and no obvious between-subject correspondence limits interpretation of the data to these three subjects. Furthermore, as this study explored only dependent variables involving patients with a particular voice pathology (vocal nodules), results cannot be generalized to other voice disorders. Subjects also did not constitute a homogeneous group. They varied in age, size of nodules and investigator's impression of severity of voice disorder.

Clinical Validity

The relevance of the results of this study to clinical practice is an important consideration. Considerable time and effort was involved in incorporating the data collection for aeromechanical and perceptual analyses into the treatment program. Equipment for the aeromechanical data collection was very specific and relatively expensive. In order for clinicians to include the aeromechanical data collection as a part of their treatment routine, there must be evidence that laryngeal airway resistance measures are clinically useful. At this time, on the basis of this study, it is not possible to state that laryngeal airway resistance is a useful indicator of progress in treatment or a reliable correlate of perceptual judgement. The factors that threaten the internal and external validity of these results also limit their clinical validity. Further well-controlled research is necessary to determine whether laryngeal airway resistance might indeed be a clinically relevant indicator of progress in treatment.

Conclusions and Suggestions for Further Research

In conclusion, results from the present study suggest that laryngeal airway resistance did not change in the same direction or to the same degree with treatment across the three subjects, nor did the unique changes that occured appear to be sensitive to changes in the status of subjects' vocal nodules with treatment. Although the nodules disappeared entirely in two subjects and resolved noticeably in the third, corresponding laryngeal airway data remained essentially unchanged, decreased moderately or increased dramatically with treatment. These results cannot be used as evidence that laryngeal airway resistance may not change reliably with voice therapy, however, because the validity of the resistance data was compromised by significant instrumental and procedural errors, and possibly by baseline instability in some or all of the subjects. In addition interpretation of results must be tempered with the recognition that sample size was small and subjects were heterogeneous in terms of age, size of nodules before therapy and severity of voice disorder. Although significant correlations occurred between data representing changes in subjects' laryngeal resistance with voice treatment and listeners' perceptions of those subjects' voices across the treatment period, the validity of the correlations is rendered ambiguous by a number of procedural, sampling and psychophysical factors. Artifacts and errors in the voice recordings, particularly with respect to loudness, may have affected the consistency of the recordings and influenced listeners' perceptions of them. The syllable samples obtained for aeromechanical and perceptual analyses were adequate for the former but probably not for the latter, while the sentence samples appear to have been more informative for perceptual analysis but were inadmissable for aeromechanical analysis. Finally, intra- and interjudge agreement data were variable and only moderately reliable which is consistent with reports in the literature for similar rating tasks and renders questionable the validity of the perceptual data used in the correlational analyses.

Future studies attempting to monitor laryngeal airway resistance could minimize possible procedural problems in the following ways. It would be very beneficial for investigators to provide subjects with considerable time to practise the task for LAR data collection and measurement before beginning to collect data. Once data collection begins, careful inspection of equipment and placement of mask and catheter on the subject will minimize problems such as those that occurred here with power supplies and mask leaks. In addition, careful inspection of the aeromechanical and acoustical data immediately following each data collection session will serve to identify any samples that are at all problematic.

A replication of this study's first research question (laryngeal airway resistance) incorporating the above guidelines and utilizing a larger number of subjects is suggested. It is recommended that potential subjects with allergies causing nasal congestion not be permitted to participate in the study in order to ensure that the voice disorder is purely a functional one.

The reliability of listeners' perceptual judgements will continue to be an unresolved difficulty in studies involving voice quality assessment. Nevertheless, investigators whose research involves perceptual judgements should ensure that a standard connected-speech sample of adequate duration is utilized for clinicians to rate and that such a sample is obtained under the best possible recording conditions. Particular consideration also should be given to choosing the perceptual dimensions for listeners to rate. In this study it is possible that listeners' perceptual judgements were largely influenced by the dimension of "breathiness". As the subjects progressed through voice therapy, they were taught to use easier voice onset. This vocal behavior may have encouraged a "breathy" voice quality, ultimately resulting in listeners' perception of the subjects' voices in treatment to be "worse" than they were in baseline. In addition, the dimensions of "breathiness" and "diplophonia" may not have been adequate to capture the dysphonia of individuals with vocal nodules. Future studies also may want to include the perceptual dimension of "tension" to more completely describe the vocal quality of hyperfunctional phonation associated with the presence of vocal nodules. The dimension of "breathiness" could continue to be used, but it would be helpful to have raters judge voice quality separately for each of these two perceptual dimensions (i.e., breathiness and tension). In this way, perceptual measures may be better able to reflect the aeromechanical components of airflow and subglottal pressure, respectively. Finally, in order to improve intra- and interlistener reliability, the use of "anchor" stimuli for the particular voice quality of interest (e.g., breathiness, or tension), and procedures that encourage consensual agreements between listeners during perceptual rating tasks may be useful.

REFERENCES

Aronson, A.E. (1980). Clinical Voice Disorders. New York: Brian C. Decker Division, Thieme-Stratton, Inc.

Barlow, D. H., Hersen, M. (1984) Single case experimental designs: Strategies for studying behavior change. Toronto, ON: Pergamon Press.

Bassich, C.J., & Ludlow, C.L. (1986). The use of perceptual methods by new clinicians for assessing voice quality. Journal of Speech and Hearing Disorders, 51, 125-133.

Boone, D.R., & McFarlane, S.C. (1988). The Voice and Voice Therapy. Englewoods Cliffs, NJ: Prentice-Hall.

Bruning, J.L., & Kintz, B.L. (1987). Computational Handbook of Statistics. Glenview, IL: HarperCollins.

Colton, R.H., & Casper, J.K. (1990). Understanding voice problems. A physiological perspective for diagnosis and treatment. Baltimore, MD: Williams & Wilkins.

Deal, R.E., & Emanuel, F.W. (1978). Some waveform and spectral features of vowel roughness. Journal of Speech and Hearing Research, 21, 250-264.

Flach, M., Schwickardi, H. & Simon, R. (1969). What influences do menstruation and pregnancy have on the trained singing voice? Folia Phoniatrica, 21, 199.

Gerratt, B. R., Kreiman, J., Antonanzas-Barroso, N. & Berke, G.S. (1993). Comparing internal and external standards in voice quality judgements. Journal of Speech and Hearing Research, 36, 14-20.

Greene, C.L., & Mathieson, L. (1991). The Voice and its Disorders. Whurr Publishers: New Jersey.

Harris, T.B. (1971). Aerodynamic and perceptual characteristics of normal and dysphonic children. Unpublished master's thesis, Purdue University.

Higgins, M.B., & Saxman, J.H. (1989). Jitter across the menstrual cycle. Journal of Voice, 3, (3) 233- 243.

Hillman, R.E., Holmberg, E.B., Perkell, J.S., Walsh, M. & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. Journal of Speech and Hearing Research, 32, 373-392.

Hoit, J.D. & Hixon, T.J. (1992). Age and Laryngeal airway resistance during vowel production in women. Journal of Speech and Hearing Research, 35, 309-313.

Isshiki, N. (1964). Regulatory mechanism of voice intensity variation. Journal of Speech and Hearing Research, 7, 17-29.

Kahane, J.C., & Beckford, N.S. (1991). The Aging Larynx and Voice. In D. N. Rapich (Ed.), Geriatric Communication Disorders (pp.165-186). Austin, TX: Pro-Ed.

Kempster, G.B., Kistler, D.J., & Hilldenbrand, J. (1991). Multidimensional scaling analysis of dysphonia in two speaker groups. Journal of Speech and Hearing Disorders, 51, 125-133.

Kostyk, B. (1996) Laryngeal airway resistance in teachers with vocal fatigue. Unpublished master's thesis, University of Alberta.

Kreiman, J., Gerratt, B.R., & Precoda, K. (1990a). Individual differences in voice quality perception. Journal of Speech and Hearing Research, 35, 512-520.

Kreiman, J., Gerratt, B.R., & Precoda, K. (1990b). Listening experience and perception of voice quality. Journal of Speech and Hearing Research, 33, 103-115.

Kreiman, J., Gerratt, B.R., Precoda, K., & Berke, G.S. (1992). Individual differences in voice quality perception. Journal of Speech and Hearing Research, 17, 153-163.

Kreiman, J., Gerratt, B.R., Kempster, G.B., Erman, A., & Berke, G.S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. Journal of Speech and Hearing Research, 36, 21- 40.

Kunze, L. (1962). An investigation of the changes in subglottal air pressure and rate of air flow accompanying changes in fundamental frequency, intensity, vowels, and voice registers in adult male speakers. Unpublished doctoral dissertation, University of Iowa.

Langhans, J. (1981). Laryngeal airway resistance during vowel production in adult females. Paper presented at the Annual Convention of the American Speech and Hearing Association, Los Angeles.

Leeper, H., & Graves, D. (1984). Consistency of laryngeal airway resistance in adult women. Journal of Communication Disorders, 17, 153-163.

Lotz, W.K., & Netsell, R. (1986). Developmental patterns of laryngeal aerodynamics for speech. Association for Research in Otolaryngology Abstracts, 134.

Lutz, R. (1986). PERCI - PC Manual. Chapel Hill, NC. Microtronics Corp.

McGlone, R.E. & Hollien, H. (1963). Vocal pitch characteristics of aged women. Journal of Speech and Hearing Research, 6, 164-170.

McReynolds, L.V., & Kearns, K.P. (1983). Single-subject experimental designs in communicative disorders. Baltimore, MD: University Park Press.

Melcon, M., Hoit, & Hixon, T. (1989). Age and laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders, 54, 282-286.

Moran, M.J., & Gilbert, H.R. (1984). Relation between voice profile ratings and aerodynamic and acoustic parameters. Journal of Communication Disorders, 17, 245-260.

Morrison, M.D., Rammage, L.A., Gilles, M.B., Pullan, C.B., & Hamish, N. (1983). Muscular tension dysphonia. Journal of Otolaryngology, 12 (5), 302-306.

Netsell, R. (1969). Sub-glottal and intraoral air pressures during the inter vocalic contrast of /t/ and /d/. Phonetica, 20, 68-73.

Netsell, R., Lotz, W., & Shaughnessy, A. (1984). Laryngeal aerodynamics associated with selected voice disorders. American Journal of Otolaryngology, 5, 397-403.

Ottenbacher, K.J. (1986). Evaluating clinical change: Strategies for occupational and physical therapists. Baltimore, MD: Williams & Wilkins.

Prater R. J., & Swift, R.W. (1984) Manual of Voice Therapy. Austin, TX: Pro-ed.

Prather, E., Hedrich, D., & Kern, C. (1975). Articulation development in children aged two to four years. Journal of Speech and Hearing Disorders, 40, 179-191.

Reed, C.G. (1980). Voice Therapy: A need for research. Journal of Speech and Hearing Disorders, 45, 157-169.

Rothenberg, M. (1982). Interpolating sub glottal pressure from oral pressure. [Editorial letter] Journal of Speech and Hearing Disorders, 47, 218-224.

Sant, L., & Logemann, J. (1970). An investigation of the variations in air flow and subglottal pressure with changes in pitch and intensity of phonation. Paper presented to the Annual Convention of the American-Speech-Language-Hearing Association, New York.

Sataloff, R.T., Lawrence, V.L., Hawkshaw, M., & Rosen, D.C. (1991). Medications and their effect on the voice. In R.T. Sataloff (Ed.), Professional Voice: The Science and Art of Clinical Care. New York: Raven Press.

Seigel, S. (1956). Nonparametric statistics. Toronto, ON: McGraw Hill.

Shaughnessy, A., Lotz, W., & Netsell, R. (1981). Laryngeal airway resistance for syllable series and word productions. Paper presented at the Annual Convention of the American Speech and Hearing Association, Los Angeles.

Shipp, T., & McGlone, R. (1971). Laryngeal dynamics associated with voice frequency change. Journal of Speech and Hearing Research, 14, 761-768.

Shipp, T., (1973). Intraoral air pressure and lip occlusion in mid vocalic stop consonant production. Journal of Phonetics, 1, 167-170.

Silverman, F.H. (1993). Research design and evaluation in Speech-Language Pathology and Audiology. Prentice Hall: Englewood Cliffs, NJ.

Smitheran, J.R., & Hixon, T.J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. Journal of Speech and Hearing Disorders, 46, 138-146.

Smitheran, J.R., & Hixon, T.J. (1982). A reply to Rothenberg. Journal of Speech and Hearing Disorders, 47, 220-224.

Stathpoulos, E.T., Hoit, J.D., Hixon, T.J., Watson., P.J., & Solomon, N.P. (1991) Respiratory and laryngeal function during whispering. Journal of Speech and Hearing Research, 34, 767.

Thompson, A.E., & Hixon, T.J. (1979). Nasal airflow during normal speech production. Cleft Palate Journal, 16, 412- 420.

Ward, K.L. (1992) Acoustic characteristics of pathologic larynges: An acoustical and perceptual study of changes in dysphonia with voice therapy. Unpublished master's thesis, University of Alberta.

Wilson, F.B. (1972) The voice-disordered child: A descriptive approach. Language Speech and Hearing Services in Schools, 4, 14-22.

Yumoto, E., Sasaki, Y., & Okamura, H. (1984). Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. Journal of Speech and Hearing Research, 27, 2-6.

APPENDIX A

SUBJECT CONSENT FORM FOR INITIAL CONTACT BY THE PRIMARY INVESTIGATOR

To whom it may concern:

In this package [Appendices B and D] you will find information about a research study involving patients with vocal nodules. If you think you are interested in participating in this study, but would like more information about the project, please sign the form below. Your signature will give permission to the person doing the study (Marnita Grams) to contact you by telephone. At that time, you will be provided with more information about the study and may ask questions regarding it. Thank you for your consideration.

____________________ Subject's Signature	____________________ Date
____________________ Subject's Printed Name	____________________ Home Telephone Number
	____________________ Other Telephone Number
____________________ Witness	

APPENDIX B

VOCAL NODULES TREATMENT STUDY

Department of Speech Pathology and Audiology
Faculty of Rehabilitation Medicine
University of Alberta, Edmonton

Principal Investigator: Marnita Grams, B.Sc.
(403) 439-8327 residence
(403) 492-5990 office

Advisor: Anne Putnam Rochet, Ph.D.

INFORMATION FOR POTENTIAL PARTICIPANTS

Purpose

This document describes a research study which is concerned with changes in voice quality and voice production associated with therapy for vocal nodules. The purposes of this study are 1) to determine the effectiveness of a measure of voice production, called laryngeal airway resistance, in monitoring changes in the voice across a period of voice therapy, and 2) to study the relationships between the measures of laryngeal airway resistance obtained and changes in voice quality associated with therapy.

The results of this study should help clinicians provide better ways of monitoring changes in voice production and voice quality during therapy for vocal nodules.

Who is Eligible to Participate?

To be eligible to participate you must 1) have vocal nodules confirmed via an examination called flexible fiberoptic video nasendoscopy performed at the Glenrose Rehabilitation Hospital's Voice Clinic, 2) have abstained from smoking for at least two years, 3) have no history of chronic allergies, chronic indigestion, respiratory illness, neurological disorders (e.g., strokes, parkinsonism), or speech and hearing problems, and

4) have abstained from taking certain medications such as anti-psychotics, anti-depressants, diuretics, antihistamines, anti-hypertensive agents, adrogenic agents and beta blockers in the last two weeks. In addition, you must pass a brief hearing screening test.

PROCEDURES BEFORE, DURING AND AFTER TREATMENT

Your participation in this study will be required over a 12- week period during which you will be seen three times per week: the 12- week time period is divided into two phases, a four- week pre-treatment monitoring period, and an eight -week period of voice treatment. Because this is a research study concerned with monitoring changes in the voice that occur with treatment, it is vital that each participant be monitored for a length of time before therapy that is at least half as long as the treatment period. As it is possible that the menstrual cycle may have an effect on the voice, female participants will be asked to keep the principal investigator informed of each onset of the menses during the course of this study.

During the first session, an interview will be conducted to confirm your eligibility for this treatment study. At that time you will be asked questions concerning the your voice problem, voice use habits, general health, and health history including illnesses that may have implication for voice, accidents involving head or neck injury, head and neck surgery or major surgery requiring intubation for general anesthesia, indigestion problems, respiratory problems, chronic allergies, neurological disorders, speech or hearing problems, smoking, and use of medications. In addition you will be asked to assess the impact of your voice problem on your occupational/ social life. You will then undergo a brief hearing screening test that lasts approximately three minutes. You must pass this test in order to be eligible to participate in this research, because the presence of a hearing problem may cause changes in voice quality or affect your ability to monitor your voice during therapy.

During the first four weeks of the research study (called the baseline phase) you will be asked to come in three times per week for approximately 10-15 minutes. During these visits audio tape recordings of your voice and measures of Laryngeal airflow and pressure during voice production will be obtained as you repeat a series of simple one-syllable utterances. To do this, you will be trained by the investigator to say the syllable "pea" seven times on one breath at a controlled rate and loudness level. Your rate will be trained with a metronome, and your loudness will be monitored with a sound-level meter.

Once you have been trained to produce the utterance "pea" with appropriate rate and loudness, tape recordings of your voice will be made at the same time as airflow and air pressure are measured. You will repeat the syllables with your face placed in an anesthesia-type mask and a tiny tube between your lips at the corner of your mouth. The mask collects the air flow; the tube senses pressure in your mouth. You will then say the syllable "pea" seven times on one breath using the rate and loudness levels you were trained to use. There are no physical risks to you by undergoing this procedure. You will be able to breathe freely while wearing the mask, and the procedure will not be painful, although you will need to press your face firmly against the mask.

After the four-week baseline phase, the eight-week treatment phase of the research study will begin. During this portion of the research study, you will be asked to come in three times per week for sessions that will last one hour. The first 15 minutes of each session will be used to record your voice, and to obtain measures of Laryngeal airflow and pressure as described in the previous paragraph. The remaining 45 minutes will be spent in voice therapy, during which you will learn new ways to use your voice that will counteract the kinds of vocal behaviors that are thought to cause vocal nodules.

When the treatment phase ends, you will be scheduled into the next available Voice Clinic at the Glenrose Rehabilitation Hospital for a second flexible fiberoptic video nasendoscopic exam to document any change in the appearance of your vocal nodules.

Appointments will be scheduled to fit your schedule and that of the principal investigator, as conveniently as possible either during the day or in the evening. A sum of $30.00 will be paid to each participant to cover parking expenses associated with the pre-treatment phase of the study.

The research will be conducted either in the facilities of the Department of Speech Pathology and Audiology in the Faculty of Rehabilitation Medicine (Corbett Hall) on the University of Alberta campus or in the Department of Communication Disorders at the Glenrose Rehabilitation Hospital . The execution of the research and the conduct of its investigation are governed by the University of Alberta's policy related to ethics in human research. In accordance with these regulations, your welfare and dignity will be protected throughout the study, and your anonymity will be ensured in the event of any publication or presentations derived from the research. The data obtained from your participation will be coded according to a system known only to the principal investigator, will be stored securely, and will be destroyed when the final reports of the research are complete. There are no liabilities to you as a participant in this research, although it will require a considerable time commitment on your part. The flexible fiberoptic video nasendoscopic exam, which you experienced as part of the diagnostic procedures at the Glenrose Rehabilitation Hospital's Voice Clinic, is a necessary follow-up procedure that will be repeated as soon as possible after treatment ends. In each case, it will be performed using a topical anesthetic to make it as comfortable for you as possible.

Your inquiring about this study in no way obligates you to participate in it. If you are eligible and do consent to take part, you are free to withdraw your participation at any time, without ill-will or jeopardy to your future health care at the Glenrose Rehabilitation Hospital, the University of Alberta or its Hospitals, or your Health Unit. Please feel free to ask any questions you might have regarding this research study. Marnita Grams is the investigator responsible for the is project. She can be reached at 439-8327 (residence) or 492-5990 (office).

Thank you for your consideration of this information.

Vocal Nodules Treatment Study
Marnita Grams, B.Sc. SLP
Department of Speech Pathology and Audiology
2-70 Corbett Hall
University of Alberta
Edmonton, Alberta
T6G 2G4.

APPENDIX C

TELEPHONE SCREENING FORM

Name:________________________________ Date of Screening:____________________

Address: ______________________________ Telephone: __________________________

D.O.B. _____________________ Age:____ Occupation: ________________________

Referral Contact:____

Date of Flexible Fiberoptic Video Nasendoscopy (FFVN) confirming presence of nodules:

____

1. When did you first notice a problem with your voice?

2. Have you ever had any speech-language or voice therapy in the past?

3. Do you smoke? Have you ever smoked ? If yes, when?

4. Do you presently suffer from respiratory illness such as a pneumonia, or asthma
or do you have chronic allergies that cause you to have nasal congestion, coughing, excessive mucous production and post nasal drip?

5. Do you suffer from chronic indigestion? Do you have heartburn frequently?

6. Have you been taking any medications over the past two weeks? (e.g., antidepressants, antipsychotics, amphetamines, antihistamines, birth control pills).

Are you currently taking any over-the-counter medications? (e.g. antihistamines and decongestants).

7. Have you ever had any neurological diseases (e.g., stroke, or Parkinson's)?

8. Would you consider participating in a research study over a twelve week period? During eight of the twelve weeks, voice therapy would be provided.

9. Do you have any questions about the study?

APPENDIX D

VOCAL NODULES TREATMENT STUDY

Department of Speech Pathology and Audiology
Faculty of Rehabilitation Medicine
University of Alberta, Edmonton

Principal Investigator: Marnita Grams, B.Sc.
(403) 439-8327 residence
(403) 492-5990 office

Advisor: Anne Putnam Rochet, Ph.D.

SUBJECT'S CONSENT FORM

Name: ______________________________

Date: ______________________________

I have read the attached information regarding the vocal nodules treatment study which will be conducted by Marnita Grams. I understand the contents of the information form, the procedures involved, the purpose of this research and the time commitment required of me. I have felt free to ask any questions regarding the study, and these have been answered to my satisfaction.

I understand that:

1. During the first session, an interview will confirm my eligibility for this treatment study. At that time I will be asked questions concerning my voice problem, voice use habits, general health and health history regarding illnesses, accidents involving head or neck injuries, head and neck surgery, indigestion problems, respiratory problems, chronic

allergies, neurological disorders, speech or hearing problems, smoking, and use of medications. As well I will be asked to assess the impact of my voice problem on my occupational and social life. This information will be recorded on my voice case history form. I will then undergo a hearing screening test which I must pass to be eligible to participate in this research.

2. This vocal nodule treatment study will span a 12- week period during which I will visit either Corbett Hall at the University of Alberta or the Glenrose Rehabilitation Hospital three times per week. Each visit during the twelve week period will involve having my voice recorded as I repeat a series of one-syllable utterances and having measures of air pressure and air flow related to my Laryngeal airway resistance sampled at the same time. I will be trained by the investigator to say the syllable "pea" seven times on one breath with appropriate rate and loudness. Once I have been trained to produce this utterance, I will say it while audio tape recordings of my voice are made and airflow and air pressure are sampled. To do this I will place my face into a mask with a tiny tube placed between my lips at the corner of my mouth. The mask collects airflow and the tube monitors air pressure in my mouth as I speak the syllables. While wearing the mask, I will be able to breathe freely and the procedure will not be painful, although I will need to press my face firmly against the mask. There are no physical risks to me by undergoing this procedure.

In the first four weeks of the study, I will be asked to come in three times per week for sessions that will last only 10-15 minutes. The visits during this period of the study will only involve having my voice recorded as I repeat a series of one-syllable utterances and having measures of my air pressure and airflow related to my Laryngeal airway resistance sampled as described above. No treatment (voice therapy) will be provided during this time. These pre-treatment recordings and measures of Laryngeal airflow and pressure are required to provide valid comparisons with the data obtained from me during the treatment period. I will be paid $30.00 to cover my parking expenses at the University of Alberta or at the Glenrose Hospital during this pre-treatment phase of this research.

During the eight week treatment period, I will come either to Corbett Hall at the University of Alberta or to the Glenrose Rehabilitation Hospital three times per week for one hour sessions. Voice recordings and measures of my air pressure and airflow related to my Laryngeal airway resistance will be obtained at the beginning of each session, followed by approximately 45-50 minutes of voice therapy.

3. The only invasive procedure I will undergo in this research study is the flexible fiberoptic video nasendoscopic examination which I have experienced as part of the diagnostic procedure for vocal nodules in the Voice Clinic at the Glenrose Rehabilitation Hospital. In order to document the effectiveness of voice therapy I will be asked to consent to undergo this procedure once more at the Voice Clinic immediately following voice therapy. A topical anesthetic will be used to make the procedure as comfortable as possible.

4. The information obtained from my participation in this study will help clinicians provide better ways of providing and monitoring treatment for individuals with voice disorders.

5. The execution of this research study and the conduct of its investigators are governed by the University of Alberta's policy related to ethics in human research. In accordance with that policy my welfare will be protected, and my anonymity will be assured in any professional publications or presentations derived from this research. The data will be coded according to a system known only to the principal investigator, will be kept confidential and stored securely, and will be destroyed when the final reports of the research are complete.

I will attempt to stay with this research study for the 12- week duration of pre-treatment and treatment periods. Nevertheless, I understand that I am free to withdraw my participation at any time without jeopardizing my future health care at the Glenrose Rehabilitation Hospital, the University of Alberta or its Hospitals or my Health Unit. My signature below affirms my voluntary consent to participate and acknowledges my receipt of a copy of this consent form.

Subject's Signature	Subject's Printed Name
	Principal Investigator's Signature
Date	Witness

APPENDIX E

VOICE CASE HISTORY

Name: ______________________ D.O.B. ____________ Age:_______

Address:__________________________ Telephone: ________________

Referred by: ____

Principal Investigator: _____________________________________

Date of Evaluation: _______________________________

Date of Evaluation at Voice Clinic: _____________________

Diagnosis at Voice Clinic: ___________________________________

History of the Voice Problem (Adapted from Boone, 1988)

1. Description of Problem:

2. Possible Cause of Problem:

3. Onset of Problem:

4. Reason for Referral:

5. Variability Through Day:

6. Voice Usage:

7. Known Abuse/Misuse:

Health History

1. General Health:

2. Physical Problems Related to Voice:

3. Illnesses and Allergies:

4. Accidents involving head and neck injuries or requiring airway intubation:

5. Surgery: (Head and neck, or requiring intubation)

6. Indigestion:

7. Neurological Disorders:

8. Medications (i.e.,, antidepressants, antihistamines, antipsychotics, birth control pills, stimulants, antihypertensive agents, etc.):

9. Smoking:

10. History of Voice Problems:

11. Previous Voice Therapy:

12. Hearing:

13. Impact of Voice Disorders on Occupational/Social Life:

Hearing Evaluation

1. Pure-tone audiometric screening

Date:____________________ Pass/Fail _____________

Behavioral Observations (i.e., throat clearing, use of intensity, pitch, vocal quality)

Confirmation of Vocal Nodules diagnosis:

Date: ____________________

APPENDIX F

Voice Treatment.

During the treatment phase of this research study, various techniques that are commonly used to treat hyperfunctional voice disorders will be used. Various tasks will be required of the subject and feedback will be provided regarding the subject's performance. Following a subject's accurate production of a particular vocal technique, positive reinforcement consisting of verbal praise will be provided (e.g. "good", nice easy onset", etc.). When the subject fails to use a technique accurately in a particular word, phrase, sentence, etc., the clinician will provide an explanation to the subject about what he or she did wrongly, and offer a demonstration of the correct production if necessary. The treatment protocol over the eight week treatment phase will consist of most or all of the following five components as applicable for each individual:

1) Explanation of the nature of vocal nodules from physiological, acoustical and aerodynamic standpoints. The clinician will illustrate these explanations with pictures of vocal nodules and will discuss the implications the nodules have on the physiology of voice production and on the quality of voice. Printed materials regarding the etiology of vocal nodules will also be provided to the clients.

2) Explanation of vocal hygiene and identification and elimination of vocally abusive behaviors: The principles of vocal hygiene will be explained by the clinician, and written material regarding instructions for good vocal hygiene will be provided to the subject. Together, the subject and clinician will identify abusive vocal behaviors used by that particular subject and discuss what can be done to prevent these. Vocally abusive behaviors will be charted and recorded daily by the patient and monitored over time. The chart will be reviewed by the clinician at the beginning of each treatment session.

3) Relaxation techniques will be used at the start of each session to reduce whole body tension. Techniques used will be as follows:

a) Quiet breathing

The patient will be taught to breathe in slowly through the nose and out through the mouth.

b)Normalizing Posture/Whole Body Relaxation.

The patient will begin the exercises standing with feet apart, and arms raised. Further instructions are as follows:

i Drop arms, feel shoulders in lower position. Repeat 2-3x.

ii Drop head to chest, roll from side to side, shifting body weight from foot to foot.

iii Add arm movement as comfortable. Keep whole body in motion.

iv. After a few minutes, add some easy sighing. Vary vocalizations from sigh, to variation of vowels, nonsense sounds. Be flexible in pitch. This is meant to relax and warm up voice. Clear sounds are not necessary.

c) Demonstration of "chewing" exercises will be used to teach the subject to use exaggerated movements of the jaw, lips and tongue. Chewing exercises are thought to indirectly reduce tension in the voice (Boone, 1988). These exercises using multisyllabic words will be used at the beginning of treatment sessions to reduce tension in the articulators and in the Laryngeal mechanism.

4) "Placement of the voice in the oral-facial mask," a technique described by Boone (1988), will be used to reduce ay tension present in the Laryngeal area, and to change the pitch to a more natural level, thereby improving vocal quality and Laryngeal efficiency. "Placement of the voice" will be introduced to the subject through humming exercises during which the s/he will be encouraged to concentrate on the sound vibrations felt across the bridge of the nose and around

the nose in the mid-facial area. A treatment hierarchy for this technique will proceed from humming to monosyllabic words containing nasal consonants, to multisyllabic words containing nasal consonants, to short nasal and oral phrases, to nasal and oral sentences, to nasal and oral paragraphs and finally to conversation. The subject will continue to "place the voice" throughout the treatment hierarchy. After each production, the clinician will judge whether the production was adequate or inadequate based on auditory judgement. A 90% success criterion will be used over 20 training trials (e.g., 20 nasal phrases) in order for the subject to move on to the next level of the treatment hierarchy.

5) Elimination of "hard glottal attack" (Boone, 1988). Subjects will be trained to develop "easy onset of voice" to eliminate abrupt voice onset through exercises progressing from isolated vowels to /h/ and vowel initiated words, phrases, sentences, to reading of paragraphs, and conversation. The Yawn-Sigh approach, as outlined by Boone (1988) may be used to elicit relaxed, easy phonation. As in technique #5, a 90% accuracy criterion will be set as a requirement for moving on to each subsequent level of the treatment hierarchy. Again, the clinician will determine whether a production was adequate or inadequate based on auditory judgement.

APPENDIX G

(Figures G-1 to G-6)

Figure G-1. Subject 1: Syllable Perceptual Ratings Data Set.

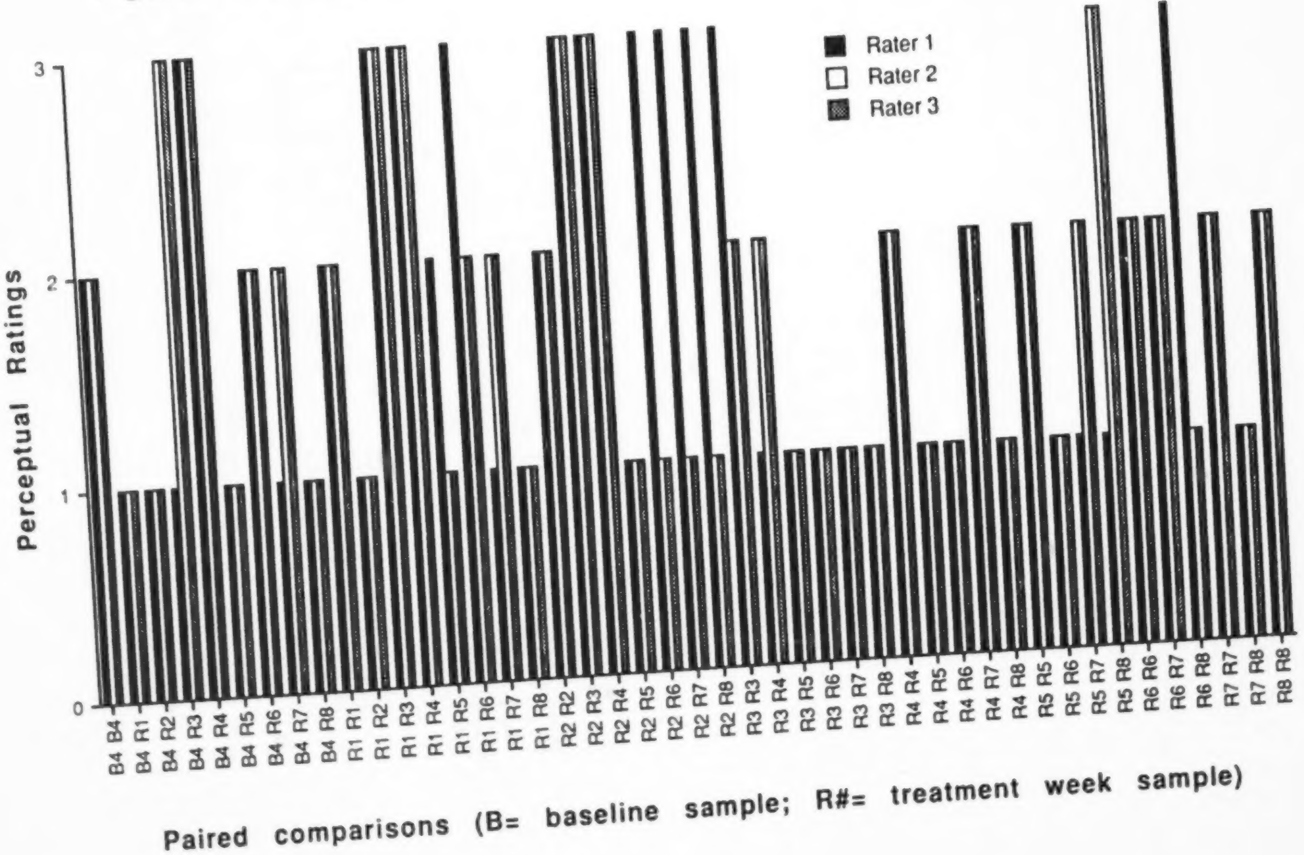


Figure G-2 Subject 1: Sentence Perceptual Ratings Data Set.

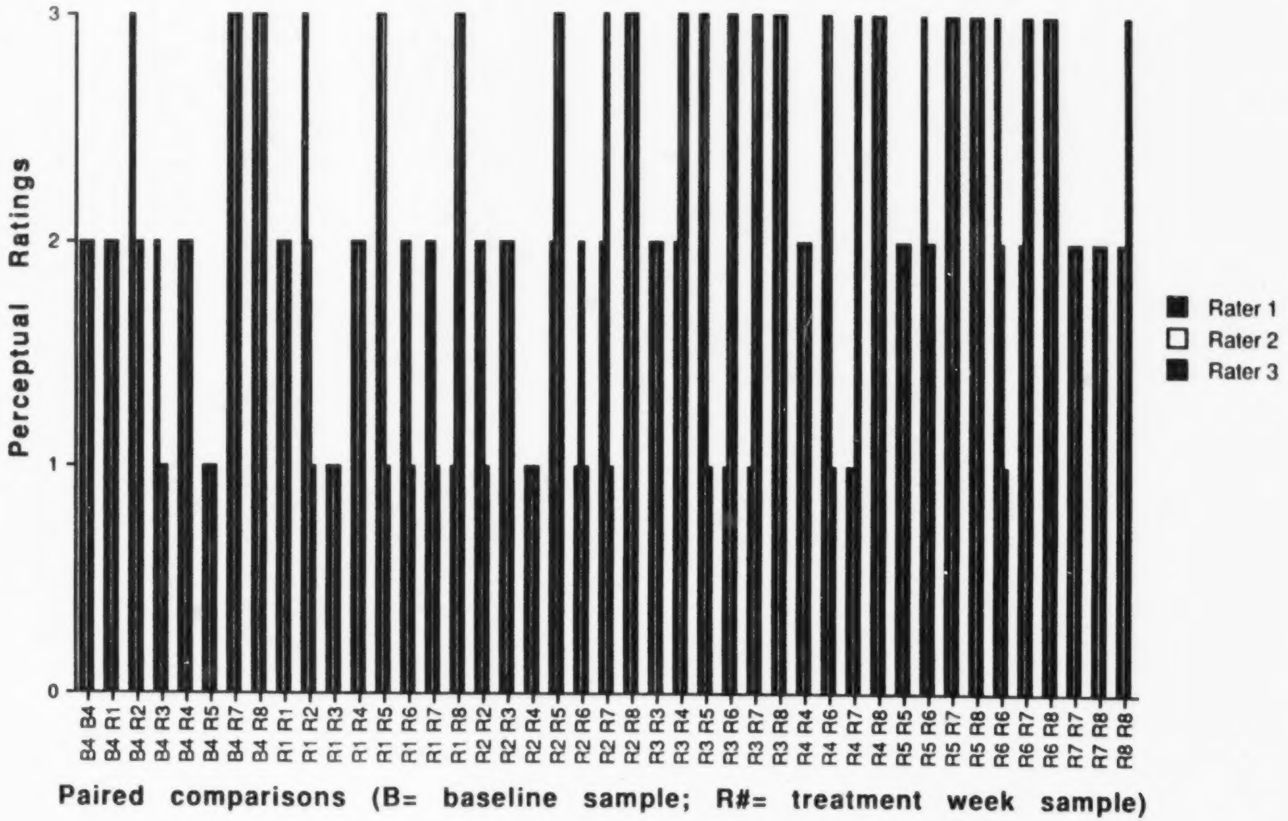


Figure G-3. Subject 2: Syllable Perceptual Ratings Data Set.

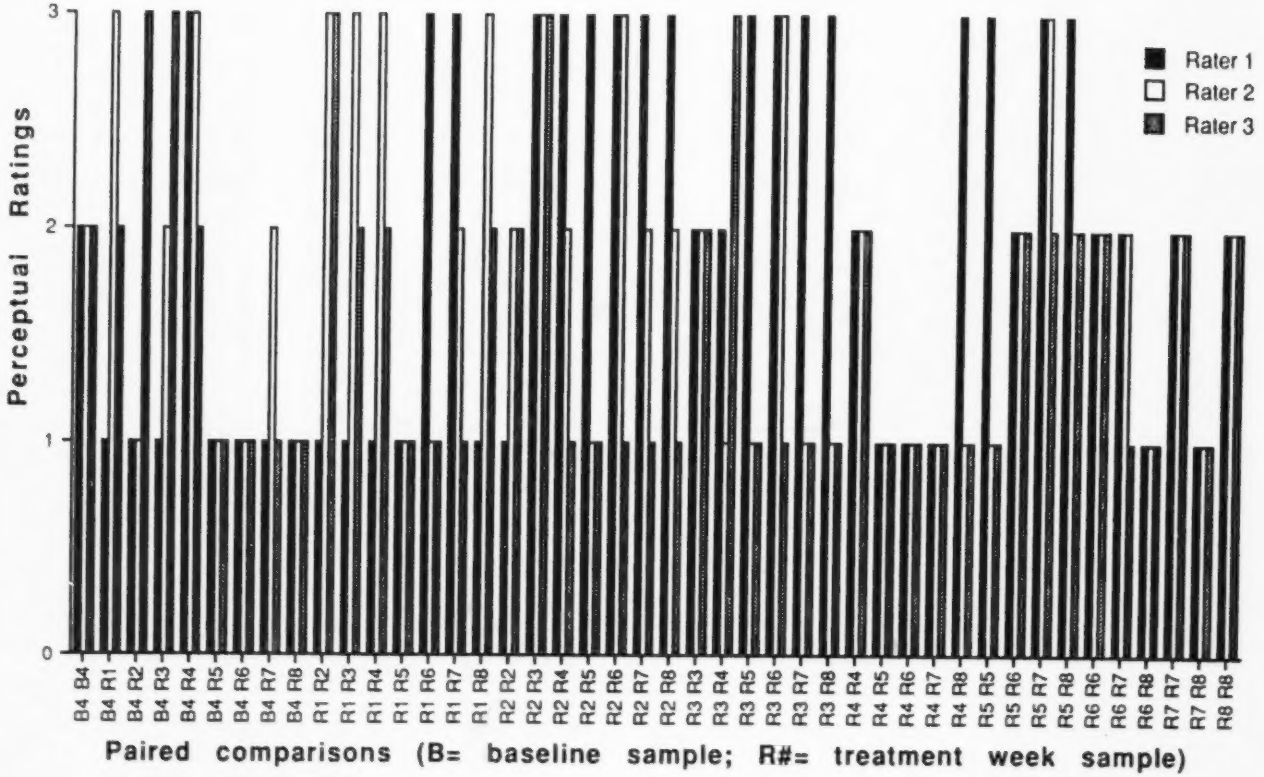


Figure G-4. Subject 2: Sentence Perceptual Ratings Data Set.

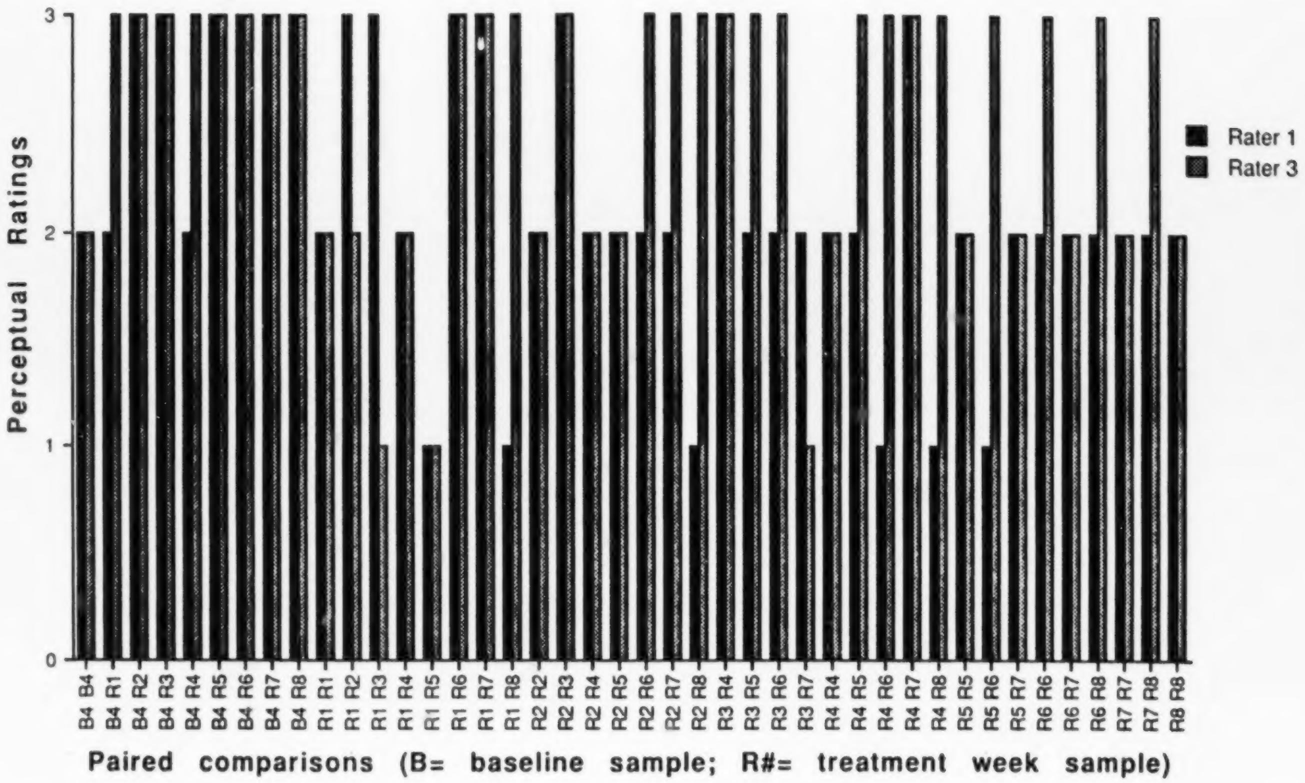
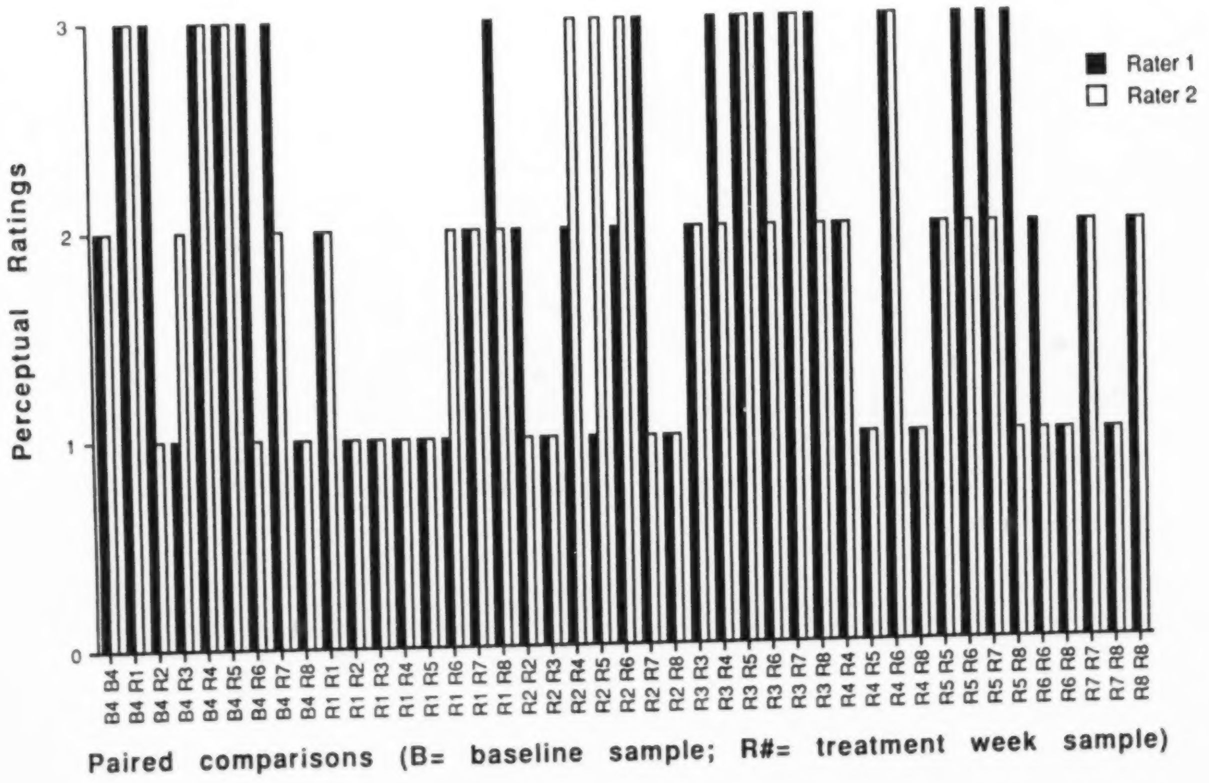
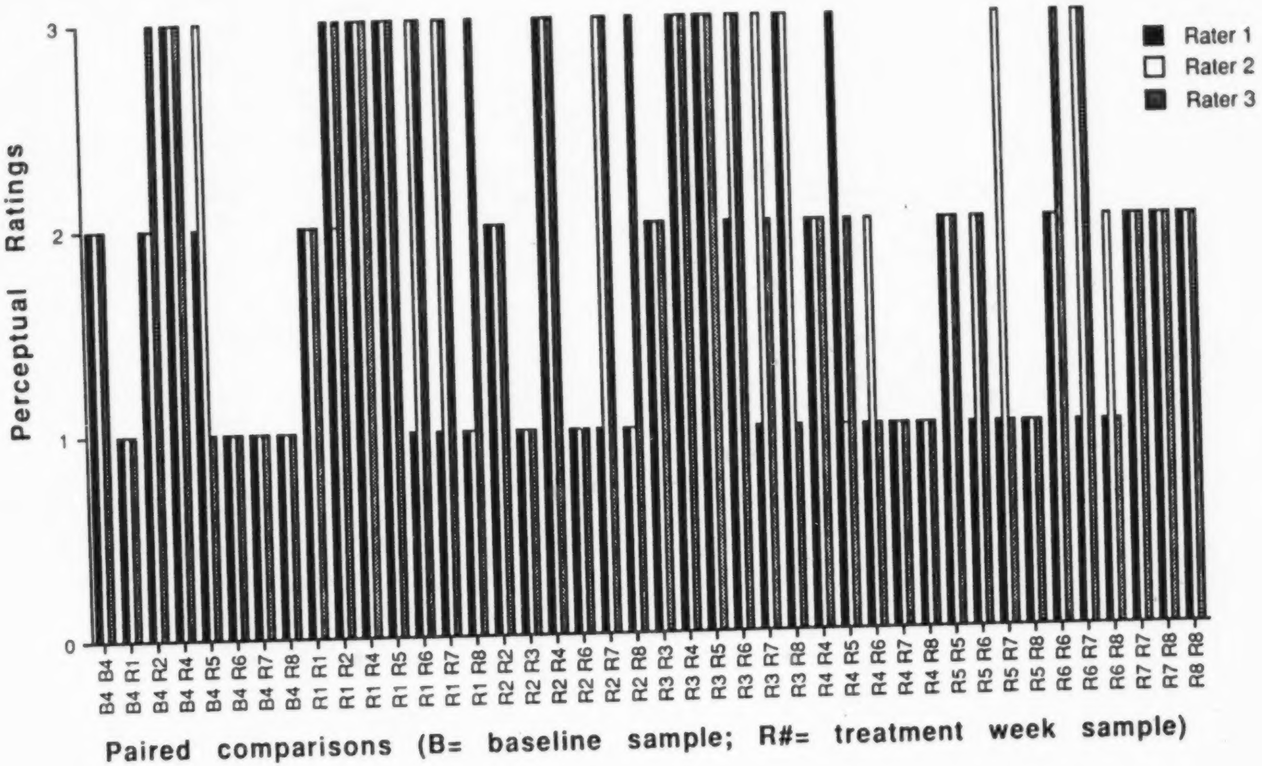


Figure G-5. Subject 3: Syllable Perceptual Ratings Data Set.



Figure G-6. Subject 3: Sentence Perceptual Ratings Data Set.